

How far can China be trusted?

Nearly forty years after Hiroshima, the United States is to conclude an agreement on nuclear cooperation with China, which has a similar deal with Japan. But the safeguards are not good enough.

"FATHER in Heaven, the Senators could use an eight-day week." Thus the invocation by the chaplain of the US Senate on Tuesday morning last week, with the continued lack of a budget for the year beginning on 1 October heavy on his mind. Now the US Congress has the new agreement with China on nuclear cooperation to worry about as well. Not for the first time, new China has faced a would-be partner with an awkward dilemma in domestic politics. Last year's negotiations about the future of Hong Kong could have turned nasty if the British had been strongly placed enough to exercise reasonable scepticism about vague Chinese promises for the future. Now it is Congress's turn to decide whether it can accept verbal promises, "non-paper agreements" as they have been called (see page 381), as substitutes for the formal physical safeguards on US uranium that would ordinarily be a feature of an agreement such as this, but which China has "adamantly refused", according to Dr Kenneth L. Adelman, director of the US Arms Control and Disarmament Agency. In the end, Congress will probably be forced to swallow its pride. But it will be a painful process.

The proposed agreement falls short of what it would have been reasonable to expect in at least three ways. First, it says nothing about the arrangements under which China will be allowed to reprocess (so as to extract plutonium from) nuclear fuel supplied from the United States. Instead, there is an agreement to agree — that lawyers' abomination — that suitable arrangements will be worked out if and when the occasion arises. Failure to agree within six months will allow China to go ahead "on an interim basis", which may be a handy pistol to hold at the heads of future negotiators from the United States. Second, although the agreement contains a promise that nuclear material will not be diverted to the manufacture of nuclear weapons, the provisions for verification consist only of a vague undertaking that visits to Chinese nuclear plants may be arranged through "diplomatic channels". Third, while the preamble to the agreement includes a reaffirmation of the principles of the charter of the International Atomic Energy Agency (IAEA), which China joined in January last year, there is nothing like a promise that China will not assist other states (Pakistan?) in the development of nuclear weapons. China has asked to be taken on trust, and the US administration has reluctantly agreed.

Congress will find this package unpalatable for a variety of reasons, partly legal but mostly political. The administration's contention that the China agreement is consistent with the Atomic Energy Act and the 1978 Nuclear Non-Proliferation Act (NNPA) is no doubt technically correct. The NNPA allows the United States to make nuclear agreements with nuclear weapons states that contain none of the safeguard arrangements to which non-nuclear weapons states must be subjected. But this let-out is permissive only. The notion that any state which has gone so far down the path of nuclear development as to be able

to make nuclear weapons for itself should be exempted from safeguards is contrary to the spirit of the NNPA and, if generally accepted, would be widely regarded as an incentive for non-weapons states to embark on making their own bombs. In any case, China now is only half a nuclear power. Its nuclear weapons are made from uranium-235, obtained from expensive diffusion plant. The proposed agreement will accelerate the development of plutonium technology in China, if only by providing experience with reactor construction. The NNPA says nothing on that score.

For what it is worth, the agreement with China now proposed by the United States goes further than those extant with two other nuclear weapons states, Britain and France, which may consider that they have a better claim than China to be trusted in the United States; US uranium finding its way into their nuclear fuel cycles is safeguarded under the cooperation agreement with Euratom. So the administration could easily find Congress demanding that it should seek a waiver from the strict provisions of the NNPA, and then be faced with the hazards embodied in Senator William Proxmire's amendment last year of the export administration act: asking for a waiver would be a humiliation, but would improve the chances of approval.

Stumbling

How, the administration will be asking itself, can it have stumbled into the position of having to carry such a can of worms to Congress? A decade ago, when China returned to the community of nations, it offered itself as a substantial market for nuclear power stations from elsewhere. Foreign reactor salesmen were crawling all over China, and it seemed only prudent to the United States that there should be a cooperation agreement with China so that US manufacturers could join in the hunt. But the passage of time and the dawning of economic realism in China have changed this prospect. The plan for a dozen nuclear power stations right away has been cut back to just three while China is understandably determined to be a designer and constructor of nuclear power stations in its own right, and thus a potential customer for reactor components, not the huge turnkey projects that make the mouths of Western contractors water. In the changed circumstances, the commercial value to the United States of the proposed agreement is probably negligible, but access to US technology is still vital to China. So, in that sellers' market, why cannot the United States insist on safeguards? Because if they were to withhold a cooperation agreement with China, its ability to function as the chief sponsor of China's re-entry into the international community would probably be seriously impaired.

The crucial question now is that of China's policy on the proliferation of nuclear weapons. The most serious impediment to a rational answer is China's touchiness about its sovereign status, which has led it to demand (and now to be offered) the widest possible latitude within the framework of the NNPA. The most serious worry is that China is not a signatory of the Non-Proliferation Treaty (NPT), and thus not bound to refrain from helping others make nuclear weapons. Yet the climate has improved since Mao Tse Tung, when China's line was that the spread of nuclear weapons would have the virtue of undermining the dominance of the then superpowers, the Soviet Union and the United States. The administration's case for the agree-

Biological manuscripts

Contributors are reminded that, with the transfer of the Biological Sciences Editor to the Washington office, it will be helpful if they will in future send **four copies of all manuscripts** offered for publication, either (as at present) to London or Washington.



ment now proposed rests on the changes that have since taken place. China has indeed joined IAEA. Chinese officials are reported to have said that Chinese collaboration with non-nuclear weapons states would be subjected to IAEA safeguards (as in the agreement with Brazil on reactor development at the end of last year). But otherwise the administration can offer only quotations from Chinese leaders' public speeches to the effect that China does not "help other countries to develop nuclear weapons". These are hardly assurances of the kind the US Congress expects from the administration's lawyers.

How might this perplexing situation be resolved? The ideal is that China should join the NPT and thus be bound by the obligation not to help others make nuclear weapons. On the face of things, there is no reason why China's present leadership should not be persuaded that such a step would insulate present policies on proliferation from whatever domestic changes may lie ahead. There is some force in the administration's claim that even an unsatisfactory agreement on nuclear cooperation will enable it to influence policy in the years ahead; it should be asked for a reasoned estimate of the chances of Chinese membership of the NPT when it comes to plead its case for the agreement in the coming months. China, of course, may argue that it is no differently placed from France, a nuclear power which has also declined to join the NPT, but that is only partly true; France has formally declared that it will behave as if it were a member, and has joined the Nuclear Suppliers Group (the "London club") to coordinate safeguard policies. That would be a half-way house for China, but is it asking too much that both these black sheep should now be persuaded, by a judicious blend of cajolery and pressure, to join up?

The harder nut to crack stems from the fact that any assistance from outside with the construction and operation of reactors will help China become a fully fledged nuclear power, with plutonium as well as uranium bombs in its arsenals. Is the United States (and other states with cooperation agreements with China — Belgium, Britain and West Germany) reconciled to that prospect? The rest of the world has not yet come to terms with the notion that China is already a nuclear power and will become an even more substantial one. Circumspectly, by seeming not to have a declaratory nuclear strategy, China has helped others to set this question aside. But if British and French nuclear weapons become a factor in the bilateral arms control talks between the Soviet Union and the United States, as they did in the earlier series aborted in 1983, should not China also be counted? And how could that be accomplished without consultation with China, perhaps even its active involvement in the Geneva negotiations? Here again the recent record is far from discouraging; Chinese participation in the UN committee on disarmament (also at Geneva) is generally recognized to have been constructive. If the United States hopes to win influence on China's nuclear policy by means of an agreement falling short of what it wants, may it not expect its influence to work on that wider canvas? That, certainly, is another issue that the US Congress should take up with the custodians of the non-paper agreements with China — before acquiescing, as it must in the end, in the administration's urgent pleas that more harm than good would be done by throwing out the proposed agreement. □

Commons debate. The opinion is widely shared, and may even be true. Certainly within the present framework, the British government's support for research is inadequate, almost calculated to engender collapse by the erosion of institutions and the spirits of those who work in them. The select committee's report is a model of clarity (and brevity) but says very little that is unfamiliar. Some of its recommendations, such as that the cost of paying research council superannuation should be a separate line item in the government's annual budget (see page 384), seem almost to have been designed as a convenience to Sir Keith Joseph, believed himself to have been converted to the view that there should be more money. Why, in these circumstances, does nothing happen?

The most arresting line in the select committee's report is that, as things are, and without more money for basic research, the "government will not achieve its own objective" of keeping British basic science vigorous. That could have been the committee's text, and it should have made more of it. The British government's attitude during the past six years towards the cost of research has been that it is for the scientific community to manage its own affairs within a budget which has, if anything, increased in real terms, although not by much. If the community then makes a hash of the job, it only has itself to blame. In exactly the same spirit might parents ask a child to clothe himself on a monthly allowance, and then say it is no fault of theirs that he has spent so much on what is called gear that he has no shoes to wear to school? One obvious respect in which the analogy is not exact is that in the management of research, the scientific community and the government have an interest in keeping research healthy: the government's wider objectives will be unattainable if the system collapses. That is why it cannot suffice for government to stick to its policy of constant budgets without caring for the consequences.

A particular and long-familiar set of issues will illustrate the problem. Three of the four science-based research councils have traditionally carried out a large proportion of their research in their own institutes, while the fourth has recently gone part of the way down that road in the belief that this is an economical way of providing academic researchers with expensive central facilities. *Nature* has long held that the arrangement robbed the councils of flexibility, but it stems from the time, at least in agriculture and medicine, when the universities collectively were not a viable alternative; more recently (in 1972), it was sanctified by the Rothschild proposals for the organization of basic science, which made in-house research institutes natural repositories of government commissions of applied research. But now, in changed circumstances, it is clear that the councils are over-committed to in-house institutes, which seem especially inflexible at times of rapid change and which may also be more expensive ways of getting work done than are university and polytechnic departments, crying out for research projects as they are. The councils are eager to take advantage of the new opportunities, but are prevented from doing so by the high cost of "restructuring", a euphemism for letting people go.

So who is to blame? The British government in its present mood would shrug its shoulders and say that the councils should not have let themselves get into this position in the first place. But how proper is it for a government to wash its hands of responsibility for a situation in which its predecessors have connived, however unreflectively? And even if it were that the present state of affairs could be called bad management by the scientific community in earlier decades, how can that knowledge absolve the present government from the responsibility (and present interest) in the preservation of a healthy science enterprise? Does one deny a spendthrift child the shoes he needs to wear to school? The reality, as the governments of France, Japan and the United States have recognized, is that this is a time in the development of technology when, paradoxically, it pays to spend more on basic science. But none of that means, as the select committee implies, that the problems of organization can be left on one side. Why not, for a change, solve two problems at once? □

Research responsibility

Another demand for more UK research funds, but the government ducks responsibility.

THE British basic science problem is becoming a bore, not merely for the vast majority of *Nature's* readers (who work elsewhere) but even for the British. The House of Commons select committee reporting last week is but the latest in a series of committees recently to have said that basic science in Britain is in danger of collapse for lack of funds. Within the past few months, the Advisory Board for the Research Councils (ABRC) has said so, as have numerous speakers in the recent House of

US nuclear trade

Ink still wet on China's nuclear pact with US

Washington

A row is brewing between the US administration and the Congress about the nuclear cooperation agreement with China, signed last week during the visit of President Li Xiannian of the People's Republic of China. The agreement, intended to permit the sale of reactor components and nuclear fuel to China, will offend those in Congress who look for a strict application of the spirit as well as the letter of the Nuclear Non-Proliferation Act (1978). But those who oppose trade with China (as distinct from Taiwan) will also use the occasion to make the administration uncomfortable.

The agreement is the first between the United States and a Communist government, and is also unusual in that it involves a nuclear-weapons state; US agreements with Britain on nuclear collaboration are a kind of precedent, although they have been subsumed for the past 15 years by the accord between the United States and Euratom (which has France as well as Britain as a member). The China pact includes no formal safeguards system, bilateral or multilateral, as required of agreements with non-nuclear-weapons states by the 1978 act. The administration said, when sending the agreement to Congress last week, that China's status as a nuclear power exempts it from requirements of this kind.

It remains to be seen whether Congress will agree. The way the wind is blowing should be clarified at the public hearing arranged for this week (Wednesday) by the Bonker subcommittee of the Foreign Policy Committee of the House of Representatives.

The China pact is important for several reasons, but not least because it is the first bilateral nuclear agreement to have been signed with an overtly Communist state. The Chinese have also made it clear that the sincerity of the US approachment with China will be judged by the administration's capacity to deliver an agreement, while Congress will regard the agreement as a test of the administration's commitment to the non-proliferation cause.

The delicacy of the situation is the chief explanation of the delay in signing the

agreement, initialled as long ago as April 1984, on President Reagan's visit to Beijing. But there have also been rumours, called intelligence reports, that the withdrawal of a Chinese group of nuclear scientists from Pakistan at the turn of the year was a precondition of last week's signing ceremony.

Under the Nuclear Non-Proliferation Act, Congress has 90 working days in which to object to the agreement, which means that the issue will be settled only in February 1986. Under the terms of the act, however, Congress has only thirty days to ask the administration to seek a waiver from the strict provisions of the law, which is why the Bonker committee has been so quick off the mark.

A committee staff member said last week that particular attention would be paid to the "non-paper agreements" between the administration and the Chinese government, many of which are classified. One of these is believed to be a memorandum prepared in Washington, shown to the Chinese by Ambassador Richard T. Kennedy, special adviser for non-proliferation affairs at the State Department, but assented to only verbally.

The agreement itself is written as between two equal partners, but is permissive only. The actual supply of reactor components and fuel by US companies will be for later commercial negotiations. The supply of "sensitive" materials and of technology for making weapons is excluded from the agreement.

One of the contentious features of the agreement is the absence of provisions for formal safeguards on nuclear materials supplied to China. The administration says, in the documents sent to Congress with the agreement, that safeguards are not required by current law because China is a nuclear weapons state. But the agreement does allow for visits to nuclear sites to be arranged "through diplomatic channels". The administration says that it will ensure that appropriate arrangements for visits are made before material is actually transferred.

The agreement goes some way to meet fears that materials supplied might be used for military purposes. Apart from the emphasis throughout the text on the peaceful objectives of the cooperation intended, there is a specific provision that materials supplied under the agreement will not "be used for any nuclear explosive device" or even for research to such an end.

The specific requirement of the law that nuclear fuel should neither be further enriched nor reprocessed without prior con-

sent has been met in a roundabout fashion. The agreement says that "neither party has any plans . . ." to deal in such a way with materials covered by the agreement, but says that requests at a later stage to reprocess nuclear fuel to extract plutonium will be considered favourably.

This clause of the proposed agreement says that there will be a period of six months during which the two governments must seek to negotiate an agreement on re-enrichment or reprocessing, and that the United States will for that period be able to insist that China does not go ahead for fear that it may "prejudge" the outcome. But after six months are up, China may proceed with its plans "on an interim basis".

With the third review conference of the non-proliferation treaty due to begin at Geneva at the end of August, it is inevitable that the agreement will be held up as a weakening of the resolve of the United States to hold back the proliferation of nuclear weapons. The administration, on the other hand, insists that the agreement is entirely in conformity with the law, but also draws attention to recent Chinese statements to the effect that nuclear proliferation is unwelcome. **John Maddox**

Plans trimmed

CHINA'S nuclear plans have been substantially scaled down from the ambitious programme widely talked of in recent years, according to documents assembled by the US administration in support of the agreement on collaboration with China.

There are three immediate projects in hand: a 300-MW power plant designed in China and being built at Qinshan, Zhejiang, a twin 900-MW reactor station (with Hong Kong participation) at Day Bay, Guangdong and a twin 1,000-MW reactor station at Sunan, on the southern bank of the Yangtze, north of Shanghai. All these are light-water reactors.

China's need for nuclear power seems largely determined by geography. Coal, the chief fuel for electricity production, is largely concentrated in the north and in central China, and although production increased by 8 per cent last year, it is not enough to meet demand.

The US Arms Control and Disarmament Agency says that China has enough uranium to support at least a modest nuclear programme, and that there has not yet been enough exploration to determine the full potential of China. China's indigenous nuclear technology is an outgrowth of cooperation with the Soviet Union, which came to an end soon after a Soviet research reactor was installed at the Beijing Institute of Atomic Energy.

According to the agency, there are two other research centres in the Shanghai Nuclear Physics Institute and an institute at Chengdu in Sichuan province that is chiefly responsible for reactor design. □

Nature in Washington

STEPHEN Budiansky, Washington Editor of *Nature* for the past year and a member of the Washington staff for the past three, has joined the Office of Technology Assessment of the US Congress on a one-year congressional fellowship. □

Eureka

France takes 'British' tack

PRESIDENT François Mitterand of France has travelled a long way since the heady days of his election in 1981, when he was prepared to see his government sink large sums of money into what now seem to have been ill-judged technological projects, particularly in the electronics industry. For despite putting FF1,000 million (£82 million) of government money towards the 1986 budget of Eureka — the French-inspired project to get European industry behind a series of joint high-tech products — M. Mitterand has put government support in second place to private funding, thus moving surprisingly close to the British position on Eureka.

This much emerges from a close reading of the text of his speech at the recent 17-delegation European meeting on Eureka in Paris (see *Nature* 25 July, p.288), a text just released along with a speech of his research and technology minister, M. Hubert Curien. M. Mitterand said that "no-one would be surprised" if he claimed that "finance is one of the keys to the success" of Eureka. But, he said, the key was not only a question of the volume of finance, "but also of its origin".

In choosing Eureka projects, "We must create throughout Europe the opportunity for attractive investments", said Mitterand. He made his appeal "to the market", adding almost as an afterthought, "but also to the support of governments". Only then did he offer his FF1,000 million, quite a paltry sum by the scale of the enterprises envisaged, and by comparison to the \$26 million proposed in the United States for the first five years of the Strategic Defense Initiative, which first crystallized French thinking on Eureka.

By contrast, M. Curien adopted a slightly more "from the top down" approach, speaking of the need "to propose certain concrete technological challenges to our industrialists". For Curien "there must be no delay, the international context makes action imperative", but he felt that already there was sufficient support behind "a certain number" of French proposals (for a list see *Nature* 11 July, p.97) for it to be "scarcely in doubt" that Eureka would soon be moving.

Curien saw the criteria for selecting projects for Eureka to be:

- That they should lead to the development of high technologies having economic or "strategic" importance.
- That they should lead to a product of "original performance" which could find a place on the market.
- That they should involve many partners — private or public — "which should provide a substantial part of the finance".
- And that they should unite many European countries.

In relation to the European Commission, these projects should be considered

as forerunners of Commission efforts to create afresh a real common market for products, and complementary to the Commission's precompetitive research exercises ESPRIT (information technology), BRITE (technology for old industries) and RACE (telecommunications), Curien said. Eureka completed and amplified the decision of the Council of Europe last September, also prompted by France, to create a variety of networks of European researchers (*Nature* 311, 197; 1984) where "much has already been done but much remains to do". Eureka is an "indispensable complement" to such "more general" research actions.

As for government finance, said Curien, "It is scarcely in doubt that we must draw on the public purse in proportion to the degree of technological risk of each project". But France already devotes "a substantial part" of its research and development budget to the technologies covered by Eureka. And hence "we are waiting for a parallel response from industry", he said.

Robert Walgate

Balancing the books

PERHAPS one reason why the French President M. François Mitterand is being a little cautious about committing huge sums of government money to his pet project Eureka (see above) is that the cupboard is bare. For now is the time that the French cabinet prepares the budget for 1986 — election year — and it seems the purse strings are about to be drawn very tight.

In current francs the government budget is expected to exceed the figure of 1 million million (£82,000 million) for the first time; but in real terms, this amounts to zero growth. Nevertheless, there will be growth in some areas and among the lucky ministers is likely to be the minister of research. The present incumbent, M. Hubert Curien, might expect to receive the modest budget increase of 4 per cent in real terms for which he asked. Education is also likely to be protected, but aid to industry, agriculture and other sectors may fall 4 per cent, and defence 2 per cent.

Curien's 4 per cent increase should be compared to nearly 6 per cent last year, and would put around another FF 1,500 million on the French civil research and development budget. President Mitterand has offered FF 1,000 million to Eureka. But whether this sum would come from Curien's increase, from a redistribution of existing sums, or as extra real money is not yet clear. Final details of the budget will not be decided until a cabinet meeting in September; it will then go to parliament in mid-October.

Robert Walgate

SDI

Physicists jib at inducements

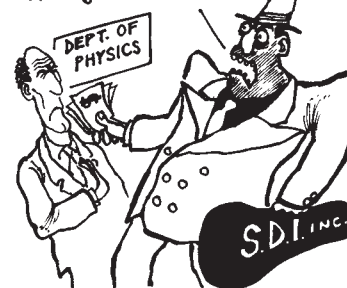
Washington

AN attempt at a boycott is being organized by several hundred academics opposed to President Reagan's Strategic Defense Initiative (SDI). They have signed a statement pledging themselves not to apply for or accept support from the SDI organization. A statement circulated by John Kogut, a theoretical physicist from the University of Illinois, condemns SDI (commonly known as star wars) as "technically dubious and politically unwise", saying it is "a step towards the type of weapons and strategy most likely to trigger a nuclear holocaust". A similar statement has been circulated by David Wright of Cornell University, and work is in progress on a joint version to be circulated to science departments all over the United States.

Kogut says he drafted his statement in response to "partisan" claims for SDI made in a brochure sent out from the SDI organization's Innovative Science and Technology Office. The brochure, which reads more like the work of a Madison Avenue advertising agency than a military department, speaks in ecstatic terms of "new and exciting" times for scientists who participate in "fascinating and challenging" SDI research.

Kogut objects to the implication that universities are "falling over themselves" to get star wars funds, and has persuaded 52 of the 72 members of his physics department to sign his protest. Physicists from

Ya mean we're makin' ya an offer ya CAN refuse?



the Institute of Theoretical Physics at Santa Barbara, from Fermilab and from the IBM Watson research laboratories at Yorktown have also signed and, according to Kogut, the number "is growing all the time". A similar petition organized earlier this year by the Union of Concerned Scientists attracted the signatures of 700 members of the National Academy of Sciences, including 54 Nobel laureates.

Kogut admits that many of those who signed his declaration are motivated by political considerations, but says that some also thought that the brochure specified unrealistic goals. The University of Illinois is a designated supercomputer

centre, and some computer scientists looked askance at the SDI office's demand for 10 million lines of error-free computer code and hardware capable of 10^{12} operations per second — one thousand times faster than anything currently available. Kogut says that many objected to large amounts of public money being spent on a proposal that would not, in their view, survive a normal scientific peer review.

So far, the protest has spread mainly within physics departments, though there are plans to coordinate it so that it reaches as many scientific departments as possible. The SDI office is maintaining an air of sublime indifference, saying that there will always be differences of opinion among experts on a complicated topic. A spokesman said the petition appeared to have had almost no effect on the response to SDI's invitation; more than 3,000 research proposals have already been received.

More than 70 per cent of the \$28 million budget for the Innovative Science and Technology Office in 1985 has been spent in US universities, and the amount is likely to rise sharply in coming years. The administration's proposed budget for star wars in 1986 is \$3,700 million, although Congress appears unlikely to grant the full request. The figures currently being discussed range from \$2,700 million to \$3,000 million.

Tim Beardsley

SDI story so far

THE SDI Innovative Science and Technology Office has so far announced seven "consortia" undertaking SDI research, involving 33 academic institutions (mostly universities), 12 private companies and sundry government laboratories. Following protests from some institutions that their names were being used without permission to lend credibility to the star wars programme, the SDI organization is now careful to say that the consortia are composed of individual scientists rather than institutions.

According to James Ionson, who heads the innovative science office, all basic research conducted on university campuses will remain unclassified. SDI research that does have to be classified is carried out in government laboratories or by private companies. The seven consortia will tackle the following subjects:

Subject	Amount (\$ million, over 3 years)
Non-nuclear space power (over 4 years)	19
Composite materials	15
Optical signal processing	9
Novel electronic and optical materials	2.5
Advanced electronics	4
Nuclear space power	6.8

Soviet research

Over-emphasis on production

DURING the past decade, Soviet science planners have put increasing emphasis on direct research contracts between the production sector and the universities and research institutes of the Academy of Sciences. Such contracts have been viewed not only as a useful source of additional income for the institutes undertaking the research, but also as an important means of mobilizing the university and academy sectors in implementing the gains of the "scientific and technical revolution" in the economy. Indeed, when addressing the special Central Committee meeting on accelerating scientific and technical progress in June, Mr Gorbachev gave strong hints that such contracts would be even more important in the future, while there would be considerable cutbacks in the (often unproductive) network of specialized research institutes belonging to the individual production ministries.

Recently, however, a leading Soviet scholar has suggested that industrial contracts are not, after all, the best way of financing university science. Dr Sergei Ambartsumyan, the rector of Erevan State University, a member of the

Armenian Academy of Science and a Deputy of the Supreme Soviet of the USSR, last month pointed out in *Izvestiya* that the contract system leads to a shortage of funds for pure research. According to Ambartsumyan, 40 per cent of Soviet scientists work in the university sector, but their research budget amounts to only five to seven per cent of the Soviet total. In order to finance fundamental research, the leading Soviet universities need at least to double their research budget by income from contracts. The production sector, however, is interested only in practical results.

Not only to the universities lack research funds; according to Ambartsumyan they are increasingly unable to attract the talented young scientists needed to carry out the planned research commitments.

Shortly before Ambartsumyan's article appeared, new pay-scales for Soviet scientists were, indeed, announced; but these seem designed to reward outstanding achievers, not to help young scientists with as yet no major achievements to their credit.

Vera Rich

Spanish research

New law to change attitudes?

Barcelona

WITH an eye to the future application of the Law of Science, already presented to the Spanish parliament (*Nature* 314, 661; 1985), a variety of scientific and institutional initiatives recently announced by the Consejo Superior de Investigaciones Científicas (CSIC) are beginning to be put into practice. According to authorities in the CSIC, which the law says is to undergo "re-foundation", the initiatives are aimed at achieving better coordination between the groups and a more dynamic administration in this institution, which accounts for much of the research done in Spain and 30 per cent of Spanish publications in international journals.

At present, the entire scientific activity of CSIC groups is organized into approximately 250 projects. These projects have been presented for evaluation to the main Spanish granting agency (CAICYT, part of the Ministry of Education and Science) and have been sent to independent referees. Several specific programmes (Programas Movilizadores) have been defined for the encouragement and coordination of research in priority topics of special scientific interest, especially those having social and economic scope. These include programmes on lasers, new materials, toxicology, food technology and natural resources.

In addition there are the thematic programmes. These are intended to establish

contacts between groups working in related subjects either in CSIC, universities or other research centres. Several of these programmes have been approved so far, including those in neurosciences, genetic engineering, space sciences, information technology and marine biology.

This may be a new way to improve contacts between CSIC institutes and the universities, traditionally strained because of distrust between Spanish state institutions. In the past two years CSIC has signed agreements with almost all Spanish universities with the intention of creating a framework that might stimulate the collaboration of research groups working in the two environments. CSIC has also revised its own institutes. In the past two years more than 80 institutes, some consisting only of one person, have disappeared, but new centres for biotechnology, for microelectronics and a Centre for Advanced Studies in Blanes, Costa Brava, have been created.

There is a general awareness that these reforms should go hand in hand with an increase in the funds for research and in staff. In fact, the budget of CSIC has been doubled in the past three years, reaching, in 1985, 21,000 million pesetas (US\$120 million). Moreover, 200 new research positions have been announced recently, in an institution employing 1,250 scientists, and a similar number will be added next year.

Pedro Puigdomènech

British research

Politicians also ask for more

THE British House of Commons Committee that serves as a watchdog for Sir Keith Joseph, Secretary of State for Education and Science, said all the right things last week about the needs of British science. But there can be no assurance that its recommendations will be in time or weighty enough to influence the budget process for the next financial year, now under way.

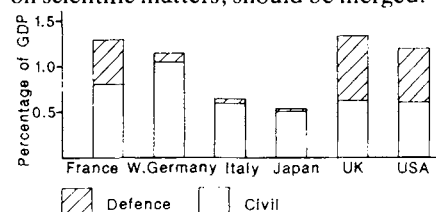
The select committee's report *The Future of the Science Budget* (HMSO) explains that the enquiry on which it is based had been begun as a "short general introduction" to the role of the Department of Education and Science in the support of research, but that the committee was persuaded, during its work, that there is substance in present concerns about the condition of British science.

Some of the committee's suggestions for change are novel, at least in the sense that they could be taken over virtually unchanged in the construction of the British government's budget for 1986 — 87. Thus the committee, noting that the total science budget, now £583 million a year, is spent largely by the research councils, suggests that the cost of paying superannuation contributions for the full-time members of their staffs should be separate from the science budget itself. The committee argues that superannuation costs are an inescapable but rising burden on

the research councils (an estimated £27 million in the current financial year).

Among the committee's other proposals are the following:

- Research councils should be compensated for the costs of closing laboratories and paying off staff if government departments cancel research commissions with council institutes, and that there should be retrospective payments to take account of the upheavals of the past few years.
- The major part of the cost of international subscriptions (as, for example, to CERN) should be borne by government departments (such as the Foreign and Commonwealth Office) rather than the research councils, on the grounds that the subscriptions are required by international obligations from which the research councils could not unilaterally withdraw.
- For the time being, the real rate of growth of the British science budget should be 3 per cent above inflation.
- The Advisory Board for the Research Councils (ABRC) and the Advisory Council on Applied Research and Development (ACARD), respectively responsible for cutting the science budget cake and for giving the British government and the public miscellaneous advice on scientific matters, should be merged.



In the last of these recommendations, the committee follows the House of Lords Committee, which reported in 1981, but does not go so far as to repeat that committee's recommendation that there should be a part-time minister for science. That, the committee says, would have required a full-scale enquiry, and "at this time we believe the main priority to be funding, not structure".

Interestingly, the House of Commons committee also shrinks from full endorsement of the idea that university budgets should be selectively determined (by the University Grants Committee), concentrating funds on universities or departments strong in research. The committee accepts Sir Keith Joseph's view that selectivity will entail "significant problems of judgement", and asks that the "dual-support system should not be undermined or weakened". But it agrees that there should be better coordination between the grants committee and the research councils on the provision of equipment.

The committee seems to have been impressed by comparisons between the state of affairs in Britain and elsewhere (see chart), although it acknowledges the diffi-

More US funds for AIDS

Washington

THE US Secretary of Health and Human Services, Mrs Margaret Heckler, has responded to congressional pressure about AIDS (acquired immune deficiency syndrome) by adding \$41 million to the budget proposed for 1986. The increase, which has been found by shifting funds from other public health programmes, will go to support community health education and studies of the epidemiology and natural history of the AIDS virus by the National Institutes of Health (NIH) and the Centers for Disease Control (CDC). The total now proposed for combatting AIDS is \$126.3 million in the year from October 1985, an increase of almost 20 per cent.

The AIDS budget has been the target of a relentless campaign by Congressman Henry Waxman, the chairman of a subcommittee in the House of Representatives with oversight responsibility for the Public Health Service. The administration's original AIDS budget proposal of \$85.5 million for next year would have meant a substantial decrease. Waxman has complained that Secretary Heckler rejected proposals by public health officials for an increased budget because of an ideological decision to freeze spending. The administration's revised proposal, which was made only after Waxman threatened to subpoena internal department documents, will now be considered by appropriations subcommittees.

New activities proposed include: a follow-up study of blood donors who are found positive by the screening test for HTLV-III (human T-lymphotropic virus type III) and studies of new HTLV-III tests, to be carried out by the Food and Drug Administration; studies of the incidence of HTLV-III in selected areas and routes of transmission, to be conducted by CDC; \$14 million for health education in co-operative agreements with individual states; and studies both of epidemiology and of possible diagnostic and therapeutic agents by the NIH.

Tim Beardsley

Auto-lobbying

THE Parliamentary and Scientific Committee of the British Parliament (Lords plus Commons) is doing what it can to make its work more pointed. A cross between a dining club and a voluntary caucus from whose meetings the press is religiously excluded, the committee has now appointed Sir Gerald Vaughan as the chairman of a ginger group that will attempt to identify important issues in the administration of science and to win a parliamentary hearing for them.

The main committee has always been a potentially valuable forum for this purpose, but the general meetings tend to follow rather than lead to important issues, and their privacy muffles their effect. Vaughan thinks that he and his small group of parliamentarians (for which Dr John Bleeby, director of a Medical Research Council unit at Carshalton, Surrey, will serve as convenor) will be able to identify emerging issues in good time. The group's first meeting, he explains, held a long discussion about continued British membership of CERN, and was persuaded that the issue deserves parliamentary attention. In this and other fields, Vaughan hopes, his group will become a lightning rod for the anxieties of the scientific community. **John Maddox**

culty of making comparisons between figures from different sources. It also points out that while the direct cost of the science budget has remained constant (or has slightly increased in recent years), indirect support, largely in the form of research commissions, has declined by roughly 10 per cent in real terms since 1981–82.

For the immediate future, the committee argues that basic research is essential for a modern industrial state, that the British government should accord it a higher priority, that academically-based research as in Britain has the particular virtue of uniting teaching and research, that the need to increase the science budget is "urgent" and that, by default, the "government will not achieve its own objective of maintaining the level of research". □

British Technology Group

UK technology transfer grows

THE British Technology Group (BTG) last week announced its annual results for NRDC (National Research Development Corporation) and NEB (National Enterprise Board). Both companies announced healthy profits, but for different reasons.

NRDC and NEB are independent organizations, coordinated by BTG, which cooperate to promote technology transfer. NRDC has two aims: to discover and initiate commercially viable projects in academic institutions, and to enter joint ventures with companies needing advice or assistance to apply technology, sharing profits in both cases. In 1984/5, NRDC received 480 applications from academic researchers, of which 301 were followed up and 149 are being actively pursued.

Until this year, NRDC had the right of first refusal on all public sector inventions, but the government has now withdrawn this privilege. Although NRDC has received slightly fewer applications as a result, the number of projects it is following up has not fallen, nor, it claims, has the standard of those it accepts.

In practice, NRDC has never enforced its right, so NRDC is not expecting great changes as a result of the ruling. However, because of potential competition from the private sector, it has set out to increase its "market awareness", publicizing itself to academic institutions, improving terms on offer to researchers, increasing the number of its liaison team and speeding up decisions on projects.

BTG does not see itself as being in direct competition with the private sector for contracts with researchers because of the "venture capital gap" (the time between setting up a project and beginning to get a return), an interval that investment companies are unlikely to provide for. BTG sees venture capital companies as potential collaborators, rather than rivals, for whichever of its projects (currently 367) develop into good commercial propositions. One obvious candidate for investment is Agricultural Genetics Ltd, in which NRDC has a 25 per cent shareholding. This company has been operating for a year and has no publicly defined goals as yet, but is likely to produce future profits in areas such as plant growth hormones.

NEB, in contrast to NRDC, is in the selling business. Since late 1983 it has been selling off all its assets, resulting in a pre-tax profit of £52 million for 1984/5 compared with £0.2 million the previous year. In the past year, 17 companies have been disposed of completely, 8 have partially gone and the remaining 56 will disappear over the next two years or so. The future of NEB after that depends entirely on the political climate in what is likely to be an election year.

Maxine Clarke

Polish universities

No freedoms next term

LAST week, the Polish Sejm (Parliament) finally passed a package of amendments to the 1982 Higher Education Act that, in the opinion of the academic community, will virtually annul the university autonomy gained in the Solidarity period. A total of 327 deputies voted for the amendments, 5 against and there were 9 abstentions.

A last-minute attempt by an independent deputy, Edmund Osmanczyk, to have the vote on the amendments deferred to the next session of the Sejm was rejected. (A general election is scheduled for 13 October). Speaking against the amendments, Osmanczyk stated that the draft tabled by the government had not taken into account the objections of the academic community as expressed by its elected representative body, the Main Council for Higher Education. Ignoring the council's opinion, said Osmanczyk, would amount to a highly immoral end to the present Sejm, and would store up trouble for the future.

Osmanczyk's complaint that the opinion of the universities had been ignored was criticized by some deputies from the ruling Polish United Worker's Party who said that the draft had been submitted to "broad public consideration". They ignored the fact that public opinion was overwhelmingly against the amendments and that the government, after formally inviting public discussion, ignored the result of that discussion. Nor did they mention the fact that the "public discussion" was conducted under somewhat difficult conditions, with the participants having to base their opinions on what were referred to as "rumours" (reports in the underground press), since the government drafts were never published in the official media.

During the Sejm debate, government spokesmen adopted the line that the amendments will not limit academic autonomy, but will simply plug a few loopholes in the 1982 act which had allowed it to be exploited by "anti-socialist elements" to "destabilize" the universities. This led to some remarkable examples of "newspeak" from the Minister of Science and Higher Education, Dr Benon Miskiewicz. According to his definition, the virtual elimination of student and junior staff participation in the university self-governance process actually makes it more democratic, since self-governance will now be exercised by senior tenured professors and assistant professors "who assume the highest responsibility for the results of research and teaching". As for the Main Council for Science and Higher Education, the 70-strong representative body elected by the universities to voice their views to the government, it does not, said Miskiewicz, represent the opinion held by the entire academic milieu but

only "individual circles and groups".

In view of the vacation, reaction from the academic community has been diffuse. The reactor of Warsaw University, Dr Kazimierz Bialkowski, has reportedly refused to comment, while one of his pro-rectors told the media, "If I say nothing, it will be bad, and if I say something it will be worse!".

Informal sampling of opinion in the main university centres suggests an atmosphere of suppressed anger that may well spread when the universities reopen in the autumn. Osmanczyk's warning of future trouble, and the warning of the Main Council, last November, that the amendments will be more likely to cause than to prevent destabilization, seem likely to prove true.

Vera Rich

Hoxha's posthumous science policy

A TWO-VOLUME collection, *On Science*, by the late leader of Albania, Enver Hoxha, has just been published in Tirana. According to the official Albanian news agency ATA, the book reflects "the correct and far-sighted policy of the Party and the theoretical scientific Marxist Leninist thought of Enver Hoxha on the development of science . . . the technical-scientific revolution in Albania, and on how to put these to the service of the revolution".

This policy envisages science and technological innovation as a major activity of the working masses (it is claimed that over 3,000 significant pieces of research were carried out last year in Albanian factories and farms), and the rapid development of mining and heavy industry (leaving agriculture and civil engineering virtually at the pick-and-shovel stage). Hoxha also proposed an intensive campaign to increase the birthrate (with a target of 3 million inhabitants of Albania by the end of the century), coupled with a major campaign to get women into industry and a programme of self-sufficiency.

Judging from the ATA reports, the book seems to have been compiled posthumously as part of a campaign of tribute which has led to the renaming, in Hoxha's honour, of virtually every important institution in the country from Tirana University downwards. Accordingly, although, in the words of ATA, the book contains "writings that are models of scientific study and work", it also contains a significant amount of previously unpublished material. Since, during his lifetime, every major speech and pronouncement of Hoxha's was published as a matter of course, this material must consist of minor pieces, included to increase the collection to the size of a major work.

Vera Rich

Calmodulin and eukaryote evolution

SIR—The ubiquity of calmodulin is primarily evidence for the ubiquity of calmodulin, *not* for evolution¹—even if found in an “ancestral type”. You say in your “Periodic extinctions undermined”² that cautious people will insist that there is as yet only circumstantial evidence that the impact caused the extinction. Would that such caution extended to biologists! Let us remind ourselves, as Schopenhauer never tires of reminding us, that one cannot make aetiological deductions from morphology.

FRANK W. COUSINS

43 Emanuel House,
18 Rochester Row,
London SW1, UK

1. *Nature* 315, 720 (1985).
2. *Nature* 315, 627 (1985).

Ragweed in China

SIR—Yeh has recently reported that ragweed hay fever is not present in China¹. This is, of course, understandable if ragweed plants are not found there, which Yeh assumed.

The Chinese botanical literature^{2,3}, however, reports that ragweed (*Ambrosia*) is not only present in China, but present in abundance — *A. artemisiifolia* in the reaches of the Changjiang (Yangtze) River and along roadsides there and *A. trifida* along village paths and riverbanks in northeastern China.

The disparity between the reported absence of ragweed hay fever and the presence of ragweed plants is puzzling, but not unique. It is analogous to the puzzle of timothy grass (*Phleum pratense*)⁴. Timothy is an important allergen in the West and, like ragweed in America, is a common cause of hay fever. Timothy is abundant along roadsides and in fields throughout China⁵, but, like ragweed, is not reported to have clinical importance there¹.

Finally, Broder *et al.*⁶, in a sample of 6,563 Americans, found 579 cases of hay fever, or a prevalence of hay fever in the United States of 8.8 per cent; but Yeh, in a sample of 6,563 Chinese, found only 226 cases, or a prevalence of hay fever in China of 3.4 per cent¹. It is curious that Yeh's sample size was identical to Broder's, but the difference in prevalence of hay fever between the United States and China may be only apparent since the studies were not concurrent and there is insufficient methodological information to permit reliable comparisons.

Even so, we are left with an important puzzle: what can explain the reported absence of hay fever due to ragweed and timothy in China when both plants are found there? Is it simply a question of ragweed and timothy sensitivities not hav-

ing been adequately searched for among the allergic population of China, or, more likely, are there genetic differences in sensitivities to various pollens between Chinese and Westerners? Finally, what role, if any, do psychological factors^{7,8} play in the prevalence of hay fever within different cultures?

WANG SIYU

Institute of Botany,
Academia Sinica, Beijing, PRC

STEPHEN A. SPONGBERG

Arnold Arboretum,
Harvard University,
Jamaica Plain, Massachusetts

JUDITH S. RUBENSTEIN

Winchester, Massachusetts

HOWARD S. RUBENSTEIN

Allergy Clinic,
University Health Services,
Harvard University,
Cambridge, Massachusetts 02139, USA

1. Yeh S.-I. *Immunol. Allerg. Pract.* 4, 23–28 (1982).
2. *Iconographia Cormophytorum Sinicorum* (Institute of Botany, Academia Sinica), 4, 488 (Science Press, Beijing, PRC, 1975).
3. Li Y., Chen Y., & Shi C. *Flora Reipublicae Popularis Sinicae*, 75, 329–330 (Science Press, Beijing, 1979).
4. Rubenstein, H.S. & Rubenstein, J.S. *J. Am. med. Ass.* 252, 3127 (1984).
5. *Iconographia Cormophytorum Sinicorum* (Institute of Botany, Academia Sinica) 5, 100 (Science Press, Beijing, 1976).
6. Broder I., Higgins M.W., Mathews K.P. & Keller J.B. *J. Allerg. Clin. Immun.* 54, 100–110 (1974).
7. Rubenstein H.S., King S.H. & London, E.L. *Adolescence* 14, 1–18 (1979).
8. Russell, M. *et al. Science* 225, 733–734 (1984).

Ageing society

SIR—In his commentary on “Prospects for an ageing population” (*Nature* 6 June, p.463), Jacob A. Brody pointed out the problems society faces with the marked increase in the number of elderly people and our present lack of understanding of the biology of the ageing process. Advances are, however, being made in the molecular biology of ageing that may ameliorate problems of ageing. An important discovery was the observed decrease in the rate of protein synthesis per ribosome with increasing age. The activity of most but not all enzymes per gram of tissue also decreases with age. The loss of activity of any individual enzyme appears to vary among individuals. Although we cannot yet reverse this decrease, we are in a position to overcome some of its consequences. One treatable consequence is the reduction in the biosynthesis of small organic molecules which act as coenzymes, neuro-transmitters, hormones, and so on; a list would include bioprotein, carnitine, choline, coenzyme Q₁₀, taurine and even purine and pyrimidine nucleotides and non-essential amino acids. If the body no longer produces sufficient of these materials, their lack must be overcome by including them in the diet. They become “geriatric vitamins”. What is needed is research in the development of routine methods for analysing these materials in serum and urine and good methods for their synthesis or isolation so they can

be used to supplement diets. Some progress has been made in this area but much needs to be done.

G. M. TENER

Department of Biochemistry,
University of British Columbia,
Vancouver, BC V6T 1W5, Canada

Falsifying Velikovsky

SIR—In his review of Henry Bauer's *Beyond Velikovsky*, Owen Gingerich observes: “Although science cannot prove that a Velikovskian scenario is impossible, it might well prove that it did not happen”¹. Although Gingerich selects Peter Huber's analysis of the Babylonian Venus tablets for this purpose, they simply are not decisive enough. Indeed, Rose and Vaughan's critique of Huber², to which Gingerich alludes, is stronger than he allows.

The best evidence proving Velikovsky's scenario did not happen is provided by Greenland's Dye 3 ice core³. This core is continuous and datable by counting annual layers back at least 7,200 years. Velikovsky's catastrophes *should* have left unequivocal markers in the ice. Not only are the expected heavy dust layers absent, but volcanic acid fallout, identified with ancient eruptions in the Velikovskian time frame, is comparable in amount to that associated with single, recent eruptions⁴. This is not what would be expected if catastrophes of the magnitude described by Velikovsky had actually happened⁵.

C. LEROY ELLENBERGER

3929A Utah Street,
St Louis, Missouri 63116, USA

1. Gingerich, O. *Nature* 314, 692–693 (1985).
2. Rose, L. E. & Vaughan, R. C. *Kronos* X:2, 1–12 (1985).
3. Dansgaard, W. *et al. Science* 218, 1273–1277 (1982).
4. Hammer, C. U. *et al. Nature* 288, 230–235 (1980).
5. Ellenberger, C. L. *Kronos* X:1, 97–102 (1984).

Novel microscopes

SIR — One variety of instrument John Maddox left out when considering “New ways with microscopes” (*Nature* 16 May, p.177) is a scanning electron microscope using heavy ions instead of electrons. This has the advantage that it slowly “sandblasts” the object, vapourizing the constituents so they can be analysed in a mass spectrometer. It also has the advantage that the structure is slowly exposed as erosion proceeds — in fact the beam can be steered to dissect the structure at molecular scale. The heavier the ion the better, for example platinum, but molecules such as phthalocyanine are a possibility.

The snag is that heavy ions need powerful magnets to focus them. It is sad to see vast sums squandered on probably superfluous accelerators and telescopes while funds for this kind of instrument are difficult to raise.

BRENNIG JAMES

Cherry Orchard, Marlow Common,
Buckinghamshire, UK

The many merits of monoclonals

Only a decade after the technique for producing monoclonal antibodies sprung serendipitously from an academic line of research, biotechnology companies are thriving on their manufacture.

IN his Nobel prize lecture (*EMBO J.* 4, 1083; 1985), Cesar Milstein describes how his interest in the mechanism by which the immune system is able to generate around 10^6 – 10^9 different antibodies from a very much smaller number of genes had reached a stage in the early 1970s at which he was convinced that much diversity was generated by frequent mutations in the antibody-producing cells. Such mutations should occur at a detectable rate in cells cultured in the laboratory. But the ordinary antibody-producing cells, B lymphocytes, cannot be cultured, so the obvious trick was to turn to the equivalent tumour cells, from myelomas. The problem then became one of finding a myeloma cell that produced a suitable antibody. Several difficulties led to a decision to fuse myeloma cells with B lymphocytes that were producing suitable antibodies, in the hope of obtaining hybrid cells that both grew well and produced antibody. The approach worked and Milstein is still using it to unravel the details of antibody production, as described in the article on page 412 of this issue. In the meantime, the technique, which was described in a paper published 10 years ago this week (Köhler, G. & Milstein, C. *Nature* 256, 495; 1975), has led to Nobel prizes for the authors and a multi-million pound industry.

Because a monoclonal antibody, the product of the technique, is produced from clones of identical hybrid cells (hybridomas), not only can it be made in unlimited quantities but it is also completely specific. Conventionally, antibodies have been produced in limited quantities by injecting an antigen into an experimental animal and later collecting its serum. Since most antigens carry various antigenic determinants, the serum will contain a mixture of antibodies, each one being produced by a different clone of B lymphocytes. A hybridoma perpetuates the production of the single antibody produced by any one clone.

For their initial demonstration of the hybridoma technique, Köhler and Milstein happened to make antibodies against sheep red blood cells, not the most useful of reagents. But, as they wrote, "It is possible to hybridize antibody-producing cells from different origins. Such cells can be grown *in vitro* in massive cultures to provide specific antibody. Such cultures could be valuable for medical and industrial use." Much to its embarrassment ever since, the National Research and De-

velopment Corporation, which until recently (see page 382) had prime responsibility for fostering UK research discoveries with an industrial potential, failed to be impressed by the commercial future of monoclonal antibodies. But Milstein was in no doubt. He therefore largely shelved his work on antibody diversity to produce a variety of valuable monoclonal antibodies, including those which are now used for all UK blood typing.

As word and the technique spread among laboratories a cottage industry began. Mostly this consisted of laboratories producing monoclonal antibodies of use for their particular project. Invariably a few other groups working in the same area could make use of the antibodies and would be provided with them on request, although occasionally with some delay or with strings attached. A few monoclonal antibodies, however, were so widely in demand that the burden on the donor became intolerable. Thus, in part, began the commercialization of monoclonals.

At the same time, others were thinking on a grander commercial scale and foresaw large and expanding markets. It was easy to identify the areas in which monoclonal antibodies would have immediate uses on the scale that warranted investment. More of a problem was how to produce them in large quantities and with high purity. Two companies have made the running. One is the UK company Celltech, which had the advantage of access to Milstein's hybridomas and expertise. The other is Damon Biotech, the US company that last week announced its involvement in a consortium that is to build a £30 million plant in Scotland to produce monoclonal antibodies — most of the finance coming from European venture capital and UK taxpayers one way or another. Celltech and Damon will produce something like 3 kilograms of monoclonal antibodies between them this year out of a worldwide total estimated to be between 10 and 15 kilograms.

Monoclonals are sold for purposes that fall into four main categories — diagnosis, immunopurification, imaging and therapy — of which only the first two constitute an established market. The diagnostic tests that use the antibodies fall into two categories. First are the assays, particularly for hormones, that already rely on antibodies but for which monoclonals have provided increased sensitivity; one benefit is the earlier diagnosis of pregnancy or ovula-

tion. Second are tests, particularly for pathogens, that have traditionally relied on several days of culturing but can now be performed in several hours with a monoclonal antibody.

For immunopurification, monoclonal antibodies offer the advantage of being able to separate one substance from a mixture of very similar molecules, not to mention the large background of unrelated substances. For example, it is only because of the availability of large quantities of monoclonal antibodies to individual interferons that several of these molecules have been purified in sufficient quantities for clinical trials.

Therapy and imaging (that is, the use of an isotopically labelled antibody to provide an image of an organ by a non-invasive means) are seen as the immediate goal for expanding companies. Already about 60 applications for such uses have been lodged with the US regulatory bodies and the bets are that at least one will be approved this year, perhaps for an antibody that inactivates the T lymphocytes responsible for the rejection of organ transplants. Another therapeutic use that has already been put to the clinical test is the removal of tumour cells from bone marrow. Further down the line is the possible use of monoclonal antibodies directed against markers on the surface of solid tumours to kill the tumours. Here the idea is either to link the antibody to a toxic drug that would be delivered specifically to the tumour or to use the antibody alone to cause the tumour to regress.

Beyond that are many thoughts, some merely fanciful, of second-generation monoclonal antibodies, most of which will be the result of engineering antibody genes. There is also a quest to produce human, rather than mouse, antibodies because, during repeated therapeutic use, mouse monoclonals may stimulate the human immune system against them.

Already the short history of monoclonal antibodies serves as a classic case with which to illustrate the economic wisdom of maintaining healthy financial support for curiosity-driven research. So widespread now is the use of monoclonal antibodies in molecular and cell biology that it seems almost inevitable that they will contribute to some of the next technologies that will spring serendipitously from research of as little apparent 'relevance' as that long pursued by Cesar Milstein.

Peter Newmark

Cognitive neuropsychology

A fruit by any other name

from John C. Marshall

ADULTS who sustain damage to the left hemisphere of the brain frequently experience gross problems in the comprehension and expression of language, shown, for example, in a study reported on page 439 of this issue¹. Expressive deficits can often be demonstrated by quite simple means. When testing neurological patients at the bedside, it is customary for the clinician to point to or hold up objects for the patient to name: a chair, a watch, a pen, the lapel of a coat or collar of a shirt, the flowers and fruit on the bedside table. The patient may grope and struggle for the correct word not only in this formal situation of naming-to-confrontation but also in spontaneous conversational speech. Sometimes a word will be uttered that is related to the intended target in sound ("hair" for chair) or meaning ("table" for chair); sometimes a circumlocution will be produced which indicates that the patient, in some sense, knows the word he is searching for ("Oh, you know, it's got four legs you sit on it") but cannot retrieve its pronunciation. In yet other patients, neologistic jargon will predominate ("That's a gair").

Such word-finding difficulty is a ubiquitous aspect of 'aphasia' that follows cerebro-vascular accident ('stroke'), brain tumour, penetrating missile injury or closed head injury. The deficit is often extreme in the acute phase of the illness, and a mild residual 'anomia' is frequently the outstanding symptom of chronic aphasia after the remission of other language difficulties. How specific or restricted can such a naming deficit be? This question has both theoretical and practical significance. The precise patterns of cognitive loss and retention seen after cerebral insult should help us discover the principles of brain organization responsible for the storage and retrieval of the vocabulary of some 75,000 words that an educated adult commands²; an accurate description of the individual patient's impaired and preserved performance is an essential prerequisite of rational remediation.

The case reported by John Hart, Rita Berndt and Alfonso Caramazza¹ suggests very considerable semantic specificity in the functional and anatomical organization of the 'internal lexicon'. Their patient, M.D., recovered well from brain infarction that initially rendered him globally aphasic for both speaking and understanding speech. The sole language impairment that persisted some 18 months post-injury was a residual anomia. But this anomia did not affect equally all classes of words; rather, it was particularly severe for the categories of fruit and vegetables. M.D. could name to visual

confrontation such comparatively rare items as an abacus and a sphinx, yet he was unable to name correctly such common fruits as peach and orange. The same deficit was also apparent on naming to tactile palpation and on naming from a verbal definition. When the patient had to sort pictures from different conceptual classes into appropriate piles, difficulty was again experienced with fruits and vegetables.

This constant pattern of category-specific impairment across different tasks might suggest that the patient has 'lost' not only the generic names of peaches and oranges, of carrots and cabbages, but also the conceptual principles whereby such natural kinds are grouped together into the superordinate categories of fruits and vegetables.

The story must, however, be more subtle than this: when the patient was given the name of a fruit or vegetable he could correctly match the word to its appropriate picture; he could similarly classify these words into their appropriate semantic categories, and could correctly judge and describe the size, colour, texture and shape of the fruits and vegetables that the words refer to. Conceptual and semantic knowledge of these categories and their exemplars thus seem to be intact. But this knowledge could only be accessed by the names of exemplars, names that the patient cannot himself elicit with normal facility. It is as if the name is the key to knowledge of fruits and vegetables; the patient cannot find the key for himself (that is, he cannot spontaneously name fruits and vegetables), neither, in the absence of the keys (that is, names) can he group exemplars into appropriate conceptual classes. But, given the name key by the examiner, the patient has no problem in unlocking his apparently intact store of information about fruits and vegetables.

Vaccines

Leprosy bacillus outwitted

from Philip Draper

AFTER more than a century of frustration at being unable to obtain sufficient material, people studying the leprosy bacillus (*Mycobacterium leprae*) and developing methods to detect and immunize against leprosy should soon have access to essentially unlimited amounts of the bacterial antigens. R.A. Young and his colleagues on page 450 of this issue describe the cloning of genes that specify five highly antigenic proteins of *M. leprae* and the production in *Escherichia coli* of fragments of the

By contrast, his direct perceptual experience of fruits and vegetables does not suffice to support conceptual categorization (or naming).

In 1770, Johann Gesner first observed that in aphasia certain classes of words were more liable to be lost than others³. Yet he was deeply suspicious of the notion that different semantic classes are located in different parts of the cerebrum, as if anatomy could mirror the spatial organization of Roget's *Thesaurus*. "The vessels of the brain", Gesner wrote, "are surely not arranged in accordance with categories of ideas, and therefore it is incomprehensible that these categories should correspond to areas of destruction." Gesner's qualms may well be justified, but, if they are, the exciting results of Hart, Berndt and Caramazza nonetheless suggest that focal damage can impair a sub-space within a thesaurus-like indexing system of names that give access to the knowledge base represented in the brain.

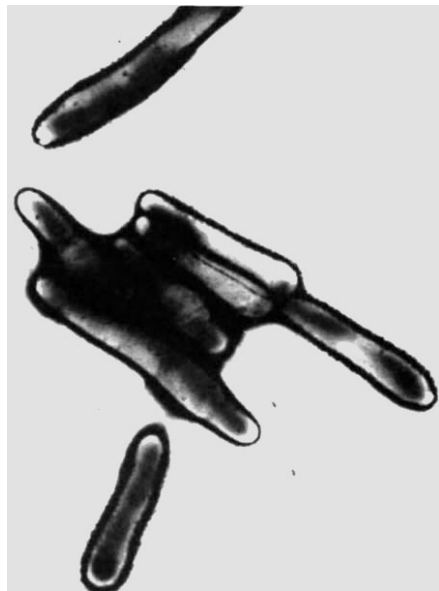
The presence of word-finding difficulties *per se* has little localizing significance; anomias have been reported after damage to many left-sided cortical and subcortical structures^{4,5}. The fact that 'naming' is not a unitary function⁶ does, however, leave open the possibility that different subcomponents of lexical processing are associated with specialized neuronal substrates. M.D.'s lesion involved the left frontal lobe of the cortex and the basal ganglia. Could a major re-entrant circuit between these structures be implicated in the semantic indexing of the mental lexicon? □

1. Hart, J., Berndt, R.S. & Caramazza, A. *Nature* **316**, 439 (1985).
2. Oldfield, R.C. *Quart. J. exp. Psychol.* **18**, 340 (1966).
3. Gesner, J.A.P. *Die Sprachamnesie* (Beck, Nördlingen, 1770).
4. Lecours, A.R., Lhermitte, F. & Bryans, B. *Aphasiology* (Baillière Tindall, London, 1983).
5. Damasio, H., Eslinger, P. & Adams, H.P. *Seminars Neurol.* **4**, 151 (1984).
6. Marshall, J.C. in *Psycholinguistics Series*, Vol. 1 (eds Morton, J. & Marshall, J.C.) 125 (Elek, London, 1977).

John C. Marshall is in the Neuropsychology Unit, Neuroscience Group, Radcliffe Infirmary, Oxford OX2 6HE, UK.

proteins that are recognized by all 13 *M. leprae*-specific monoclonal antibodies that have been tested.

This encouraging result is the outcome of a programme begun over a decade ago by the IMMLEP (Immunology of Leprosy) Steering Committee of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases. The object of IMMLEP is to produce a vaccine against leprosy, in pursuit of which it has sponsored the production



M. leprae bacilli negatively stained with uranyl acetate ($\times 12,000$).

of *M. leprae* in nine-banded armadillos — the only species, apart from man, in which this organism grows freely. Methods have been worked out for purifying the bacilli from armadillos and IMMLEP is able to offer *M. leprae* to collaborating investigators. It has been shown that killed leprosy bacilli protect mice against infection with *M. leprae* and human trials are in progress (see *Nature News* 18 July p. 183).

The possibility of applying the techniques of gene cloning to *M. leprae* was extensively discussed in 1977 at a meeting organized in Boston by Roy Curtiss III and the late Charles Shepard. Success, it was agreed, might make available candidate vaccine antigens and reduce the dependence on armadillos. DNA was soon cloned into the readily grown *E. coli* but its expression proved to be a problem; apparently the signals used by *M. leprae* to control transcription into RNA and translation into protein are not recognized by *E. coli*. Few new proteins were produced (Clark-Curtiss, J.E. *et al.* *J. Bact.* **161**, 1093; 1985).

The new method of Young *et al.* succeeds by using signals belonging to *E. coli*. Fragments of DNA from *M. leprae* are spliced on to the signal sequences and part of the *E. coli* gene for the inducible enzyme β -galactosidase, so that when *E. coli* is treated with an inducer of the enzymes, a 'fusion protein', which starts with a fragment of the enzyme and continues with the amino-acid sequence specified by the DNA from *M. leprae*, is formed.

Things are not quite so simple. Random fragments of DNA are not necessarily in the correct reading frame; displacement of the reading frame so that translation begins at the second or third base in a triplet will either halt translation or produce a nonsensical polypeptide. The solution is to ensure that fragments beginning at all three reading frames are present and then to select the desired product.

A highly efficient selection procedure is a feature of the new method. The hybrid gene is incorporated into a bacteriophage able to multiply in and lyse *E. coli*, and the fusion protein is released during lysis. A very large population of recombinant bacteriophages is produced containing sufficient different fragments of *M. leprae* DNA that all reading frames of the whole genome are represented. It is practicable to screen about a million lytic plaques, produced by spreading infected *E. coli* on to agar plates, at one time. Peroxidase-labelled antibody detects those plaques producing *M. leprae* antigens, and the antigens can be prepared in large amounts from the appropriate plaques.

The antigens are needed for several purposes but especially as reagents to measure sensitization (the production of an immune response) and its persistence by candidate vaccines. This will give some preliminary indication of activity sooner than the 10 to 15 years needed to demonstrate actual protection against leprosy. Similar reagents would be of immense value in detecting people who have had contact with *M. leprae*, either without developing the disease (apparently common) or before clinical signs appear.

Production of vaccine from *M. leprae*

antigens is a more remote possibility, though it will probably become the method of choice because production of the bacterium itself in armadillos will not be able to satisfy demand. Acquired resistance to leprosy is controlled by cell-mediated immunity, rather than antibody-mediated immunity. Research on just which antigens of *M. leprae* are recognized by T-lymphocytes in protective immunity is only now beginning, so it is not yet possible to apply Young's method to producing them. Nor is it yet known how a polypeptide antigen should be delivered to induce cell-mediated immunity.

The new method has already been applied to *Mycobacterium tuberculosis* (Young, R.A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 2583; 1985), the recent apparent failure of the traditional BCG vaccine against tuberculosis in South India having stimulated a search for new vaccines. The authors point out that it can also be applied to organisms with much bigger genomes, including eukaryotic parasites, for which immunological reagents and vaccines are badly needed. $\square$

Philip Draper is at the National Institute of Medical Research, Mill Hill, London NW7 1AA, UK.

Solid-state physics

Long-range ordering in stable semiconductor alloys

from P. M. Petroff

IT HAS BEEN known for some time¹ that stable semiconductors in the $\text{Ga}_{1-x}\text{Al}_x\text{As}$ alloy system can be produced under special conditions. This is odd because the alternating layers of AlAs and GaAs that are built up to form such compounds are meant to be miscible at the temperature required for successful growth². In a recent paper, Kuan *et al.*³ report the first observations of long range order in such semiconductor alloys. This raises important questions concerning the equilibrium phases of this system and may perhaps

explain the remarkable stability of the artificially ordered analogues.

The long range order has been found in $\text{Ga}_{1-x}\text{Al}_x\text{As}$ (where x is between 0.5 and 0.75) epitaxial layers grown on the (011) and (100) oriented cubic surface of GaAs crystals. These layers were produced by co-deposition of Ga, Al and As at high temperatures (650–750°C) using metal-organic chemical vapour deposition and molecular beam epitaxy (MBE). The transmission electron diffraction pattern of a $\text{Ga}_{0.25}\text{Al}_{0.75}\text{As}$ alloy clearly indicates

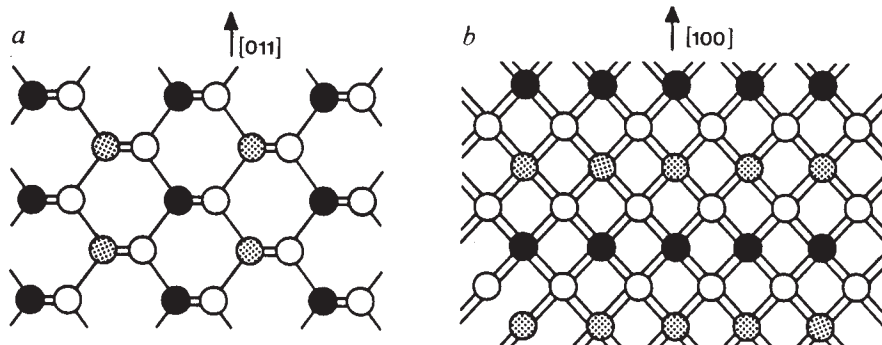


Fig. 1 Schematic representations of the crystal structure for a $(\text{GaAs})_1 - (\text{AlAs})_1$ alternate monolayer crystal with *a*, growth axis along the [011] direction and the structure projected on a (011) plane, and *b*, a growth axis along the [100] direction and the structure projected on a (001) plane. The two structures are identical and correspond to that of the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ alloy with perfect long range order. Ga is shown as black circles, Al stippled and As open.

the presence of a long range ordered phase with $\text{Ga}_{0.5}\text{Al}_{0.5}\text{As}$ composition modulation (Fig. 1a). This ordered alloy structure is the same as that of an artificially layered alloy comprising alternate monolayers of (GaAs)₁ and (AlAs)₁ on a (100) crystal surface (Fig. 1b).

This artificially ordered compound consisting of alternate monolayers of (GaAs)₁ and (AlAs)₁ can be produced¹ by MBE by exposing the (100) GaAs surface at high temperature to fluxes of Al and As, first for a time sufficient to grow an AlAs monolayer and then to fluxes of Ga and As to form a GaAs monolayer. Repetition of this deposition sequence a few thousand times produces an ordered layered compound, but only within a narrow range of substrate temperatures. Indeed, the long range order in this case is artificially produced under conditions promoting layer by layer growth².

Kuan *et al.* interpret their new results, together with the known stability of the artificially ordered compound against interdiffusion during deposition at roughly 580°C, as a sign that the $\text{Ga}_{0.5}\text{Al}_{0.5}\text{As}$ ordered structure is the equilibrium state of $\text{Ga}_{0.5}\text{Al}_{0.5}\text{As}$. More recently, other ordered alternate monolayer semiconductor systems such as (AlSb)₁-(GaSb)₁ (ref. 5) and (InAs)₁-(GaAs)₁ (ref. 6), have been fabricated by artificial layering using MBE and metal-organic chemical vapour deposition respectively. From the long range order detected in these alloys, it seems that the $\text{Al}_{0.5}\text{Ga}_{0.5}\text{Sb}$ and $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ ordered alloys are also stable compounds.

The atomic processes that are responsible for the phase separation into GaAs- and AlAs-rich regions during the co-deposition of Ga, Al and As are not understood. They probably involve extensive surface diffusion and exchange reactions between Ga and Al atoms in the upper most layers. If the ordered compound indeed corresponds to an equilibrium phase, a number of other experimental observations will need to be reinterpreted. For example, the impossibility of growing the ordered $\text{Ga}_{0.5}\text{Al}_{0.5}\text{As}$ compound by MBE on the (100) and (111) polar planes or by liquid-phase epitaxy on the polar and non-polar (110) GaAs planes⁷, suggests that an orientation dependent surface phase diagram may be required for this alloy system. Exploring the optical and transport properties of these new semiconductor alloys may increase our understanding of alloy clustering effects. □

1. Gossard, A. C., Petroff, P. M., Wiegmann, W., Dingle, R. & Savage, A. *Appl. Phys. Lett.* **29**, 323 (1976).
2. Stringfellow, G. B. *J. Phys. Chem. Solids*, **33**, 665 (1972).
3. Kuan, T. S., Kuech, T. F., Wang, W. I. & Wilkie, E. I. *Phys. Rev. Lett.* **54**, 201 (1985).
4. Petroff, P. M., Gossard, A. C., Wiegmann, W. & Savage, A. *J. Cryst. Growth*, **44**, 5 (1978).
5. Hirayama, Y., Ohmori, Y. & Okamoto, H. *Jap. J. appl. Phys.* **23**, 488 (1984).
6. Fukui, T. & Saito, H. *Jap. J. appl. Phys.* **23**, 521 (1984).
7. Petroff, P. M., Cho, A. Y., Reinhart, F. K., Gossard, A. C. & Wiegmann, W. *Phys. Rev. Lett.* **48**, 170 (1982).

P. M. Petroff is at AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA.

Frontiers of chaos

ALTHOUGH the ideas that dominate non-linear dynamics — solitons and chaos — date back to Scott Russell and Poincaré, nonlinear dynamics has flowered only since the widespread introduction of electronic computers into applied mathematics. Numerical experiments have led to the rediscovery and popularization of non-linear phenomena; and the effect on non-linear dynamics of widely available microcomputers with excellent graphics promises to be as dramatic as was the introduction of computers themselves. This can be sensed from *Frontiers of Chaos*, an exhibition that is now touring art galleries and academic institutes in Britain and the United States*, and will move to France early in 1986. The exhibition, presented under the auspices of the Goethe Institut, consists of computer graphics from the Dynamical Systems Graphical Laboratory of Heinz-Otto Peitgen at Bremen.

The most striking series of graphics deal with a Julia set (Douady, A. & Hubbard, J. H. *CRAS Paris*, **294**, 123; 1982), and the Mandelbrot set (Mandelbrot, B. B. *Physica* **7D**, 224; 1983). If z and c are complex numbers, then $z_{n+1} = z_n^2 + c$ is an iterative quadratic mapping in the complex plane. Repeated iteration of this mapping for a value of the parameter c from an initial point z_0 might lead to z approaching a fixed point attractor. For example, for $|z_0| < 1$ and $c = 0$, $z = 0$ is an attractor, or z might approach infinity. These two cases are separated by a boundary that can have the characteristics of a fractal — such a

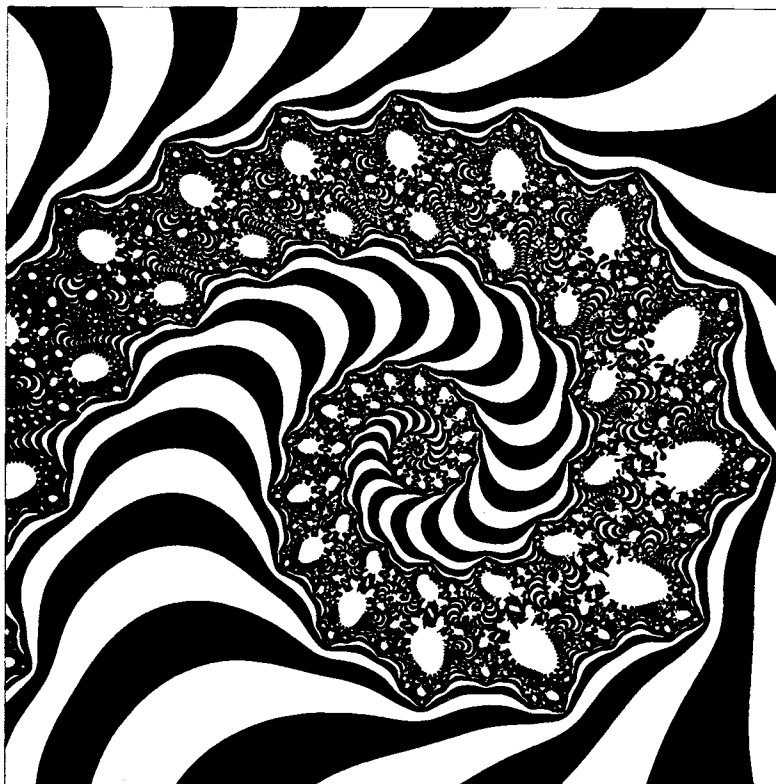
self-similar boundary is a Julia set (Peitgen, H.-O., Saupe, D. & Hasler, F. V. *Maths Intelligencer* **6**, 11; 1984). The Mandelbrot set is a subset of the set of c s for which the Julia set is connected. For the quadratic mapping it can be found as the set of c s for which repeated iteration of the mapping for $z_0 = 0$ does not give z converging to infinity.

Such a Mandelbrot set may be displayed using a microcomputer, using the video-screen as the complex plane representing c , and identifying all points on the screen where a reasonable number of iterations does not take the point $z_0 = 0$ beyond some reasonably large number, its size depending on your patience: a few hundred suffices. The picture may be elegant and will show the self-similar nature of the Mandelbrot set, with buds on buds on buds.

Introducing colour can lead to beauty: points in c outside the Mandelbrot set converge to infinity at different rates. When colour is used to code the rate, or the number of iterations it takes for the point to pass a set distance from z_0 , a wealth of self-similar dendrites, whorls and spirals emerge.

Whether these images are art, or little more than designs, is irrelevant: their beauty is incontestable. Perhaps more important, they represent a fusion of mathematics and computer graphics that can be used to model and analyse spatially complicated phenomena. **Arun Holden**

Arun Holden is at the Centre for Nonlinear Studies, Leeds University, Leeds LS2 9N2, UK.



A figure from the *Frontiers of Chaos* exhibition catalogue (reproduced with permission).

*Venues and dates from Goethe Institute, 50 Princes Gate, London SW7 or Goethe House, 1014 Fifth Avenue, New York 10028.

Membranes

A variety of calcium channels

from Harald Reuter

MANY cellular functions depend on changes in the free cytosolic Ca^{2+} ion concentration. There are two basic mechanisms by which the cytosolic Ca^{2+} concentration can be increased: by releasing Ca^{2+} ions from intracellular storage sites, and by moving Ca^{2+} ions from the extracellular space into the cell through the surface membrane. The most important structures controlling the second mechanism are proteins embedded in the lipid bilayer of the membrane, which function as 'channels' or 'pores' through which Ca^{2+} ions can move. The molecular structure of these channels not only determines which ions are allowed to pass through, that is, their ion selectivity, but also the membrane potential-dependent opening and closing kinetics, commonly called the gating of the channels. Recently it has become apparent that there is a whole variety of potential-dependent Ca^{2+} channels, each with distinct properties. On pages 440 and 443 of this issue R. W. Tsien's group at Yale reports on three types of Ca^{2+} channel in cultured sensory neurones of the chick dorsal root ganglion¹ and two types in mammalian cardiac cells².

How can one distinguish between these different potential-dependent Ca^{2+} channels? The most advanced technique is patch-clamp recording³, by which functional properties of individual channels, such as gating, conductance and pharmacological sensitivities, can be analyzed. Using this electrophysiological technique on sensory neurones, Nowicky *et al.*¹ find three unitary conductances with different kinetic features, which can be attributed to the opening of three distinct Ca^{2+} channels. The most common, which has been observed in many cells, is the L-type; re-

peated opening of these produce a long-lasting inward Ca^{2+} current through the membrane. The T-type channel, first described by Carbone and Lux⁴, opens at much more negative membrane potentials than the L-type and produces a transient inward membrane current. Both types are also present in cardiac cells⁵. The third type of Ca^{2+} channel found in sensory neurones, the N-type¹, can only be activated by large depolarizations from a very negative membrane potential. In addition to the different potential ranges at which they are gated, the three types of channels can be distinguished by different sensitivities to pharmacological interventions.

Why are Ca^{2+} channels interesting? The most important reason is their significance for cell function. The opening of Ca^{2+} channels during excitation is an essential step in the rhythmic firing of nerve cells, in neurosecretion and in the activation of contraction in cardiac and smooth muscles. Furthermore, the activity of Ca^{2+} channels can be modulated in various ways. L-type Ca^{2+} channels in cardiac cell membranes were the first potential-dependent ion channels shown to be influenced by neuro-transmitters, such as catecholamines and acetylcholine. This modulation occurs through a cascade of events and is presumably linked to a cyclic AMP-dependent phosphorylation reaction of the channel protein or of a protein closely associated with the channel⁶. Modulation of Ca^{2+} -channel activity, in turn, greatly influences the functional state of the heart.

In sensory neurones, although Ca^{2+} -channel activity is also affected by cyclic nucleotides⁶, noradrenaline exerts its effect without their involvement⁷. More-

over, L-type, but not T- or N-type, channels are influenced by 1,4-dihydropyridines², a class of drugs with wide therapeutical applications. For example, nifedipine is a Ca^{2+} -channel blocker widely used in antihypertensive therapy.

It has been a long-standing puzzle why these Ca^{2+} -channel blockers have prominent effects on Ca^{2+} channels in cardiac and smooth muscle but not in the central nervous system. It is likely that very different proportions of Ca^{2+} channels are sensitive and insensitive to dihydropyridines in these tissues. More L-type Ca^{2+} channels may open during excitation in cardiac and smooth muscle cells than in neurones, where the other channel types may prevail. It will be of considerable interest to investigate the respective densities and pharmacological sensitivities of various types of Ca^{2+} channels in different tissues in greater detail, which will also help to elucidate the functional significance of the various types of Ca^{2+} channel.

Another important observation made by Nilius *et al.*² is that T-type Ca^{2+} channels retain their functional activity in isolated membrane patches, whereas L-type channels lose theirs⁵. Metabolic input seems to be required for proper functioning of the L-type channels⁵. The maintenance of T-type channel activity in isolated membranes may explain why it has been at all possible to reconstitute Ca^{2+} channels from rat brain in planar lipid bilayers⁸.

What is known about the molecular structure of Ca^{2+} channels? Preliminary evidence from biochemical analysis of the 1,4-dihydropyridine binding site^{9,10} indicates that one of the channels may be a glycoprotein of approximately 210,000 relative molecular mass, consisting of three subunits. It is not clear whether these subunits represent the complete Ca^{2+} channel; reconstitution of all individual subunits in a lipid bilayer is required to answer this question. Further work will require gene cloning and protein sequencing to show the degree of similarity between the various types of Ca^{2+} channels. Mutants of unicellular eukaryotes carrying Ca^{2+} channels¹¹ are another promising tool in the molecular analysis of these widely distributed channel types. There is a long way to go to obtain a complete picture of the various types of Ca^{2+} channels and their respective functions. □

100 years ago

INTERNATIONAL INVENTIONS EXHIBITION

SELF-ACTING or automatic machinery has made wonderful strides of late years, and its progress in the special department of watch-making cannot be more advantageously studied than in the beautiful display of machine tools now exhibited at South Kensington. The machine tools are all labelled, and can readily be identified.

A screw-making machine. This machine is engaged in producing watch jewel screws; the size of the screws may be appreciated when we state that it takes more than 8500 to weigh one ounce troy. Lengths of wire are transformed into these tiny screws in the following manner. The machine is fed with the wire through a hollow mandrel, the wire is seized and rotated rapidly, a movable cutter is brought against it, and immediately the body of the screw is turned. Two dies are at hand which attach themselves, and they cut the thread; on reaching the limit of their cut, they pull out the wire a distance, the thickness of the screw head, for hitherto the wire has only projected the

length of the body of the screw through the mandrel head. The dies disengage themselves, and a second cutter cuts off the screw at its junction with the mandrel head. There is an alternating arm, the most conspicuous part of the machine: this takes possession of the screw as it is cut off, and, carrying it to a different part of the machine, holds the head against a small circular mill, where the notch is cut. The screw is now finished, and is discharged into a magazine by a kind of ramrod. The machine turns out 4000 screws a day.

A scape wheel tooth-cutting engine. The scape wheels to the number of sixty are threaded upon a kind of split spindle, which passes through the spaces around their arms, and holds them firmly. The spindle with the wheels around it looks like a solid rod of brass, and the cutter acts transversely so as to scoop a groove through all the sixty at once. The whole sixty wheels are cut in about twenty minutes. From *Nature* 32 296, 30 July 1885

1. Nowicky, M.C., Fox, A.P. & Tsien, R.W. *Nature* **316**, 440 (1985).
2. Nilius, B., Hess, P., Lansman, J.B. & Tsien, R.W. *Nature* **316**, 443 (1985).
3. Hamill, O.P. *et al.* *Pflügers Arch.* **391**, 85 (1981).
4. Carbone, E. & Lux, H.D. *Nature* **310**, 501 (1984).
5. Reuter, H. *Nature* **301**, 569 (1983).
6. Fedulova, S.A., Kostyuk, P.G. & Vesclosky, N.S. *Brain Res.* **214**, 210 (1981).
7. Forscher, P. & Oxford, G.J. *gen. Physiol.* **85**, 743 (1985).
8. Nelson, M.T., French, R.J. & Krueger, B.K. *Nature* **308**, 77 (1984).
9. Curtis, B. & Catterall, W.J. *biol. Chem.* **258**, 7280 (1983).
10. Borsotto, M. *et al.* *Biochem. biophys. Res. Commun.* **122**, 1357 (1984).
11. Haga, N. *et al.* *Cell* **39**, 71 (1984).

Harald Reuter is Professor of Pharmacology, University of Berne, 3010 Berne, Switzerland.

Earth sciences

Origin of mantle turbulence

from Peter Gambles

ALTHOUGH a seismic discontinuity 670 km beneath our feet is generally regarded as the boundary between the upper and lower mantle, the fact is that the very existence of 'upper' and 'lower' parts of the Earth's mantle is not proven, to the embarrassment of those concerned with earth sciences on the grandest scale.

Essentially, there are two unanswered questions — does the mantle convect as a single unit or as two discrete systems and what is the nature of the source regions that create the geochemical heterogeneities of the crust? A number of papers at the most recent European Union of Geosciences meeting* cut across traditional barriers between geochemistry and geophysics not to answer these questions, but to pose a third — what is the relationship between the two?

Taking two-layer convection as a starting point, C.J. Allègre (Institut de Physique du Globe, Paris) discussed a simple box-model consisting only of a continental crust, a correspondingly depleted upper mantle and an undepleted lower mantle. By considering geochemical and isotope observations on continental crust, ocean floor basalts and mantle xenoliths, together with the requirements of mass balance, a system of 35 simultaneous equations can be produced. Assuming a gaussian distribution of the errors and a least squares solution, one of the results that emerges is a mean age for the continents of 2.013 ± 169 Myr. This can be used to relate the present-day concentration of various elements in the crust to their initial concentration in the upper mantle before it was depleted by extraction of the crust. Comparison with the enrichments seen in mid-ocean ridge basalts (MORB) indicates that the model of extraction of the crust from now-depleted mantle is internally consistent. A few elements, such as Rb and Pb that plot off the general trend are notably volatile, while others like Re and Os are strongly siderophile and can be expected to have been drawn down into the core.

A major limitation of the model put forward by Allègre is the omission of ocean island basalts (OIB). These voluminous outpourings (forming features such as the Hawaiian islands) are isotopically distinct from normal MORB and are relatively enriched in incompatible trace elements, so that their overall character is more primitive. As OIB are commonly associated with hotspots (areas of increased heat flow), the standard explanation has been that MORB is extracted from a depleted upper mantle and OIB from primitive lower mantle. Unfortun-

ately this elegant model cannot cope with the heterogeneities observed within and between suites of OIB. Therefore various multi-component systems have been suggested¹. Explanations involving very long-lived regional heterogeneities seem to be ruled out by calculations which show that the mantle would be homogenized by mechanical turnover in less than one thousand million years². In a recent paper³, Allègre and Turcotte propose that subduction not just of oceanic crust, but also of some continental material (perhaps sediments rather than continental crust) down to a boundary layer presumably at 670-km depth, can account for the natural variations in composition. The mean age of the continents is roughly the same as the time required to evolve the observed Pb isotopic difference between OIB and MORB. Instabilities in this boundary layer could then give rise to plumes that would ultimately appear at the surface as hotspots.

But can subduction actually create a chemical boundary layer at 670-km depth? A. E. Ringwood (Australian National University), argued that it could, because of the balance between body forces acting on a subducting slab. The cold oceanic plate going down a subduction zone is more dense than the surrounding pyrolitic mantle and so exerts a pull on the unsubducted portion of the plate, causing subduction to continue. As the slab descends, it warms up and starts to disintegrate; the depleted pyrolitic lithosphere, from which the crust has been extracted, is neutrally buoyant and thus detaches and is resorbed into the upper mantle. Between about 500 and 650 km, harzburgite (the main component, together with basalt, of oceanic crust) is roughly 500 kg m^{-3} more dense than mantle material, but above and below these depths it is less dense. Thus when the melange of basalt and harzburgite reaches 650 km, it experiences no body forces, the rate of descent of the bottom end of the slab slows dramatically and a 'blob' forms. Over a period of perhaps 100 Myr, a large volume of subducted material would gather in this blob, with a root extending to depths of 900–1000 km. Geophysical evidence for this is just beginning to emerge, said Ringwood, with the detection of deep gravity anomalies at the ends of subducting plates and seismic indications of density heterogeneities.

As the blob, which is initially some 300°C cooler than its surroundings, comes into thermal equilibrium, the basalt will begin to melt and react with the harzburgite, enriching the latter in heavy isotopes. The two will also separate out because of their different densities, the basalt sinking

into the lower mantle whilst the harzburgite flows out to form a thick unstable layer at the 670-km discontinuity. Ringwood speculated that large plumes rising from this could initiate continental break-up, whilst the smallest may be trapped in the upper mantle, eventually giving rise to local geochemical heterogeneities in OIB.

D. J. Stevenson (California Institute of Technology) put forward a possible mechanism to explain the observed geochemical variations — 'magmons'. These are stable solitary waves of magma with some of the properties of solitons. Of particular interest is that magmons are capable of collecting partial melt into a packet, and that separate magmons can pass through each other with their forms more or less unchanged, but with some exchange of material. For example, a rising plume of relatively pristine lower mantle could intersect with and become contaminated by an older blob with the isotopic characteristics of continental crust. Calculations by D. R. Scott (California Institute of Technology) shows that magmons give rise to an episodicity in volcanism and can only survive in regions where the whole system is rising at much less than three centimetres per year. This excludes the mid-ocean ridges, but probably allows magmon preservation in the middle of ocean plates. Just what happens mechanically when a mantle plume is overridden by a spreading centre is not yet clear.

This leads on to a longstanding problem in igneous petrology: the geochemistry of many intrusions can be modelled by the extraction of small amounts (usually much less than 10 per cent) of partial melt from a postulated source rock, but the mechanism by which such small quantities of liquid can be efficiently squeezed out of a porous solid has remained obscure.

Armoured with assumptions, not the least of which were the requirements of chemical and textural equilibrium between the melt and residual grains, D. P. McKenzie (University of Cambridge) showed that however low the porosity, the partial melt remains physically continuous and is theoretically extractable. This allows him to model the gradual compaction of the residue and consequent expulsion of the melt fraction. Interestingly the compaction length and expulsion time turn out to be independent of density and gravity, with variations in melt viscosity being the most important factor. The separation velocities for extremely viscous melts such as dry granites are vanishingly low, but those for picritic basalts are sufficiently high for them to separate out from a 60-km thick layer in about 60,000 years — a plausible time in geological terms. McKenzie showed that tholeiitic basalts could be formed and extracted from only 0.3% partial melting of their mantle source and that melting of as little as 0.1% could lead to the rapid expulsion of volatile rich magmas, which could extract trace elements that would otherwise

*European Union of Geosciences, Strasbourg, 1–4 April 1985

have been present in later melt fractions. The model also supports the idea of very early degassing of the Earth's mantle to form the primitive atmosphere.

Studies of the distributions of noble gases lead to the conclusion that the Earth was quite efficiently degassed very soon after accretion, perhaps 4,400 Myr ago. This implies that the mantle was differentiated into enriched and depleted parts (not necessarily layers) at a very early stage. Since early Archaean tonalite (the main component of Archaean continental crust) is enriched in most trace elements, and tonalite can readily be derived from the mantle by partial melting, there would seem to be little problem except, as K. C. Condie (New Mexico Institute of Mining Technology) pointed out, the Nd isotope data require a depleted mantle source.

His solution is to postulate a variant on the plate tectonic theme. Although the ultimate driving force for tectonic movements is convection in the mantle, the forces are currently modified to such an extent by the large and rigid continents that the relationship between present-day plate boundaries and the convection cells is obscured. But the Archaean oceanic crust, according to Condie, although formed at ocean ridges much as today, then floated like a scum and gathered at the down-welling edges of the convection cells, unrestrained by rigid continents. There, the pieces of depleted basaltic crust (derived from depleted mantle) stacked up in the same way that they now do on lava lakes. By about 3,000 Myr ago, the roots of these stacks would have extended deep enough for a phase transition to take place yielding denser eclogite. At the same time, the breakdown of amphibole would have released fluids enriched in trace elements, which would contaminate the overlying crust. Finally, the enriched basaltic crust would have been dragged down to depths where extensive partial melting would occur, causing production of tonalites and the rapid growth of stable continents 3,000–2,700 Myr ago.

Returning to the subject of convection cells, McKenzie presented preliminary results of three-dimensional numerical calculations of convection at high Rayleigh numbers (the Rayleigh number defines the mode of convection; current estimates for the mantle are of the order of one million). Running the model up to Rayleigh numbers of 70,000 required over 60 hours on a Cray computer and, although McKenzie would obviously like to go much higher, some intriguing results have emerged. Under conditions of fixed viscosity and variable heat flux, the convection cells have a square planform. The zones of up and down welling are initially symmetrical, but at Rayleigh numbers greater than 50,000 this symmetry is broken. Interestingly, neither the geoid nor ocean bathymetry appears to be sensitive to the thermal structure, but the

model predicts a significant effect on the gravity field immediately above the cell. That this effect has not been detected is perhaps due to the masking effect of the thick plate above.

Although some of the connections are tenuous, the presentations at Strasbourg suggest that our knowledge of the Earth's mantle is now sufficient for geochemists and geophysicists to begin to make predictions that can be tested against evidence from both fields. The challenge for the next decade is to integrate these studies

with new data from high-pressure experimental petrology to yield a fuller understanding of the processes that determine the distribution of the continents on which we live. □

1. Zindler, A., Jaegou, E. & Goldstein, S. *Nature* **298**, 519 (1982).
2. Olsen, P., Yuen, D. A. & Balsiger, D. J. *geophys. Res.* **89**, 425 (1984).
3. Allègre, C. J. & Turcotte, D. L. *Geophys. Res. Lett.* **12**, 207 (1985).

Peter Gambles is an Assistant Editor of Nature

Cell motility

Mitotic spindles in isolation

from Peter J. Hollenbeck

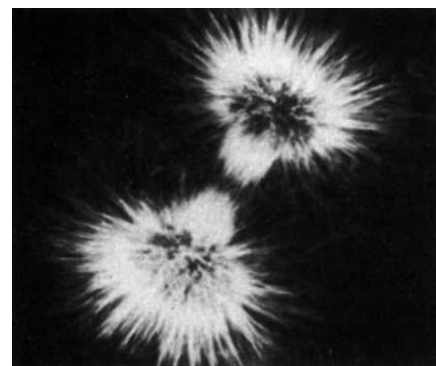
OVER 30 years have passed since Mazia and Dan¹ first isolated the mitotic spindle, liberating this organelle once and for all from the realm of artefact and establishing its existence as a coherent physical structure. In the intervening years the hapless spindle has been subjected to a barrage of experimental approaches, but has never divulged much about how it accomplishes the transport of the two chromatids that make up each chromosome to opposite daughter nuclei in the anaphase of cell division. The mechanochemistry of other motile systems such as muscle and cilia have been substantially unravelled by use of a tool that the mitosis field has lacked: a simple, functioning *in vitro* model. But a recent report in *Nature*² raises the hope that the function of the mitotic spindle too will soon be understood at a similar level, for its partial reactivation *in vitro* has been accomplished.

Cande and McDonald² report that mitotic spindles prepared from diatoms elongate in the presence of ATP. The vast majority of spindles in these preparations undergo elongation, and their associated structural changes support the notion that sliding of microtubules accompanies the movement.

As a motility machine and subject for study, the mitotic spindle is distressingly complex. It combines several types of movement, possibly of different functional origins and heaped on one another by evolution, in an organelle that is entirely dismantled during each cell cycle^{3,4}. Each chromosome is separated into its component chromatids by a movement toward the spindle poles, designated anaphase A. At the same time (or before, or after, or not at all, depending upon which organism you have selected) the spindle elongates as the poles separate in the process called anaphase B. Given the fact that the microtubule scaffolding of the spindle is assembling and disassembling during these movements, allow for some shimmying of the whole structure in preparation for nuclear migration in some species,

it is no wonder that 30 years' effort has produced, with two possible exceptions^{5,6}, no functioning isolated spindles.

So, what has led to the recent success? Certainly, the choice of diatoms as the starting material was important. Diatom spindles exhibit a structure as nearly tailored for analysis as one could imagine, as morphological studies by Pickett-Heaps and others have shown⁷. Unlike the flocculent baskets of microtubules seen in



Microtubules of the anaphase mitotic spindle in a fertilized sea urchin egg visualized by anti-tubulin immunofluorescence $\times 480$.

many organisms, the spindles of diatoms are nearly crystalline in arrangement and so they might be expected best to resist the rigors of isolation. Furthermore, they display a spatial separation of machinery that offers the hope of untangling spindle functions; microtubules associated with anaphase B form a dense central spindle distinct from the microtubules involved in chromosome-to-pole movement⁸.

It is this highly ordered central spindle that Cande and McDonald obtain in their preparations and its elongation that they have succeeded in reactivating *in vitro* — presumably reproducing anaphase B. In the presence of ATP, the zone of overlap between the two halves of the central spindle decreases in length as the spindle elongates by a similar amount. This result would seem to favour one of the plethora of outstanding hypotheses about anaphase B, by which interactions between microtubules from opposite half-spindles

in the zone of overlap generate force. Microtubule polymerization is not providing the force here, as no excess tubulin is present. (Indeed, elongation is limited in this isolated system to the length of the original zone of overlap; *in vivo*, microtubule growth during anaphase allows these spindles to elongate two-to-threefold.) The absence of structures beyond the spindle poles in these preparations also seems to exclude mechanisms by which the poles interact with elements outside the spindle to generate force. And mechanisms involving force generation by release of tension, coupled to chromosome-to-pole movement, are unlikely to explain the isolated spindle, which lacks anaphase A movements. But forces generated by interactions between microtubules from opposite half-spindles could produce the elongation in Cande and McDonald's preparations, as well as explaining the difference in extent of movement between *in vitro* and *in vivo* spindles; as microtubule growth does not accompany elongation, the *in vitro* spindle simply walks itself apart.

What more can be learned about mitosis from an isolated spindle? The wish list is long indeed. A simple *in vitro* reactivated model should allow detailed and coordinated physiological, morphological and biochemical studies of how spindles move. What is the motor that drives the spindle poles apart, and where is it located? How is the poleward movement of the chromosomes similar or different? How are microtubule growth and disassembly coupled to movement? What signals does the spindle receive from the cytoplasm to regulate its assembly, movements and disassembly, and what proteins are their targets?

Obstacles to the full exploitation of the isolated diatom spindle remain. The *in vitro* elongation is difficult to observe while it occurs because of its extreme sensitivity to light (an interesting observation in itself, and more difficult to explain in isolates than in living cells). The kinetochore microtubules must be preserved and identified in these spindles if anaphase A is to be investigated. And the problem of obtaining sufficient quantities of pure spindles for biochemical studies is clearly of primary importance. If these difficulties can be overcome, the isolated diatom spindle may prove to be for mitosis what the demembrated axoneme has been for ciliary beat or the glycerinated myofibril for muscle contraction—a key to the mysteries of a complex machine. □

1. Mazia, D. & Dan, K. *Proc. natn. Acad. Sci. U.S.A.* **38**, 826 (1952).

2. Cande, W.Z. & McDonald, K.L. *Nature* **316**, 168 (1985).

3. Inoue, S. *J. Cell Biol.* **91**, 131s (1981).

4. Kubai, D.F. *Int. Rev. Cytol.* **43**, 167 (1975).

5. Goode, D. & Roth, L.E. *Expl. Cell Res.* **58**, 343 (1969).

6. Sakai, H. *Int. Rev. Cytol.* **55**, 23 (1978).

7. Pickett-Heaps, J.D. & Tiptit, D.H. *Cell* **14**, 455 (1978).

8. McDonald, K.L., Pickett-Heaps, J.D., McIntosh, J.R. & Tiptit, D.H. *J. Cell Biol.* **74**, 377 (1977).

Peter J. Hollenbeck is in the MRC Cell Biophysics Unit, London WC2B 5RL, UK.

Gene control

Eukaryotic topoisomerases come into the limelight

from Geoffrey North

PERHAPS the outstanding structural feature of DNA is that it consists of two polymeric strands, one twined around the other. This has two important consequences: one is that it creates a topological problem for processes such as replication, where the two strands of DNA have to be separated, and the other is that it permits a topological dimension to DNA heterogeneity, which cells can potentially exploit to make DNA of identical sequence behave differently under different circumstances. This is the domain of the topoisomerases, enzymes that catalyze alterations to the topological state of DNA, and which clearly invite the attentions of those who seek structural solutions to questions about gene regulation in eukaryotes. Eukaryotic topoisomerases have been somewhat shadowy entities by comparison with their prokaryotic counterparts^{1,2}, but the evidence of a recent meeting* and a number of papers is that they are now coming into the light. The intriguing possibility has been raised that they are responsible for establishing the 'open' domains of chromatin that seem to be prerequisite for gene expression in eukaryotes. And a topoisomerase has turned out to be a major structural component of the chromosome 'scaffolds' that are implicated in defining these chromatin domains.

Yeast

The emergence of the eukaryotic topoisomerases from the shadows owes most to their genetic analysis in yeast, which has revealed functional homologies between eukaryotic and bacterial topoisomerases. The best known of all topoisomerases, bacterial gyrase, alters the state of DNA by inducing torsional stress in the reverse sense to that of the DNA helix. This process, generally referred to as negative supercoiling, depends on the ability of gyrase to make transient breaks in both strands of its DNA substrate, which defines it as a class II topoisomerase. Last year, three groups reported that budding^{3,4} and fission⁵ yeasts have single genes (*TOP 2*) for enzymes with this ability. Yeast with *top 2* mutations die during mitosis, unable to terminate DNA replication and fruitlessly trying to segregate their hopelessly entangled daughter chromosomes. But *TOP 2* seems inessential at other times in the yeast cell-cycle, for if the cell-cycles of *top 2* mutants are arrested with, for example, α -mating

factor, they can survive indefinitely⁶. So, like gyrase, yeast topoisomerase II seems to be a solution to the topological problem posed by DNA replication — though it is still not known whether it can negatively supercoil DNA.

Yeast topoisomerase I, on the other hand, seems inessential at any time in the cell-cycle. J. C. Wang (Harvard) and R. Sternglanz⁷ (SUNY, Stony Brook) reported that whilst budding yeast has a perfectly respectable gene (*TOP 1*) for a class I topoisomerase (which make only single-strand cuts in DNA) they have been unable to observe any phenotypic effect of *top 1* mutations. A requirement for topoisomerase I is, however, revealed by combining *top 1* mutations with temperature-sensitive *top 2* mutations: at the non-permissive temperature the synthesis of both DNA and RNA is affected in these double-mutant cells, which die even if their cell cycle is arrested outside mitosis.

The implication of these results is that at least some topoisomerase activity is required throughout the yeast cell cycle, but that it can be provided by either topoisomerase I or II. This requirement may relate to the observation of J. A. Huberman and colleagues (Roswell Park Memorial Institute) that supercoiled plasmids are relaxed in yeast so long as one of the *TOP 1* or *TOP 2* genes is functional, which implies that topologically constrained DNA in yeast will be maintained in a relaxed state. This observation and the lack of a requirement in yeast for topoisomerase II other than at mitosis counsel caution in interpreting the recent studies of topoisomerases in 'higher' eukaryotes that I shall now describe.

Beyond yeast

These studies fall into three main categories: observations on the effects of inhibiting topoisomerase II; attempts to identify specific sites at which topoisomerases interact with DNA; and the use of antibodies to localize topoisomerases in nuclei and on chromosomes. Experiments of the first two kinds were reported by P. Schedl (Princeton). When cultured *Drosophila* cells are subjected to 'heat shock' a series of specific and reproducible changes in gene activity ensue. Several genes are induced, with concomitant alterations to their chromatin structure. According to Schedl, this response is blocked by the prior addition of novobiocin, an inhibitor of bacterial gyrase and *Drosophila* topoisomerase II⁸. If novobiocin is added after induction, the tran-

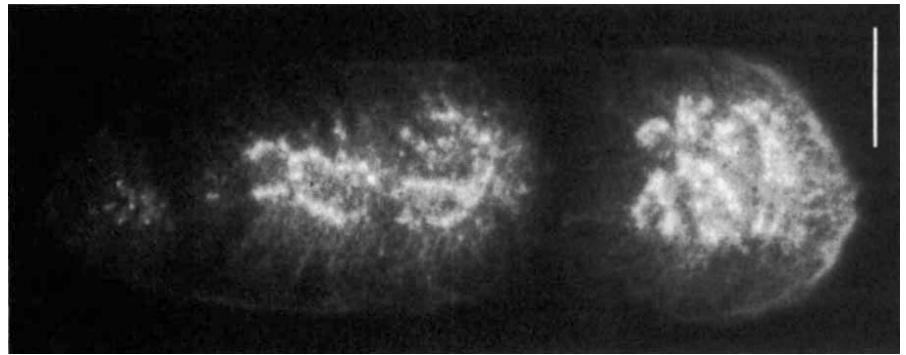
*The 1985 Cold Spring Harbor meeting on chromatin structure and expression was held at Cold Spring Harbor, New York from May 8–12 1985.

scription of heat-shock genes is inhibited, but the 'active' structure of the chromatin is maintained. So it would seem that in *Drosophila*, unlike yeast, topoisomerase II is continuously required for transcription, as well as to effect the changes in chromatin structure that accompany gene activation and inactivation.

The sites where topoisomerase II interact with DNA can be identified by blocking its reaction at a point where the DNA has been cut but not yet re-ligated, and then mapping the sites of DNA cleavage. For the divergent pair of *Drosophila hsp 70* heat-shock genes, Schedl's group has mapped the sites where purified topoisomerase II cleaves naked DNA *in vitro*⁹, and the cleavage-sites induced *in vivo* by topoisomerase II inhibitors. In both cases the sites flank the *hsp 70* genes, and most correspond to sequences previously shown to be hypersensitive to nuclease digestion. Similarly, B. Villeponteau (University of California, Los Angeles) reported that topoisomerase II cleavage sites map at nuclease-hypersensitive sites in the *Drosophila* histone gene repeat unit. And in a recent paper in *Cell*¹⁰, Yang *et al.* report that topoisomerase II cleaves within a nuclease-hypersensitive region of the SV40 chromosome when SV40-infected monkey cells are treated with a topoisomerase II-inhibiting antitumour drug. These results are intriguing in view of the known association of hypersensitive sites with gene activity¹¹, though whether they are functionally significant, rather than an artefact of the greater accessibility to enzymes of DNA at these sites, remains to be seen.

Open chromatin

As well as having discrete specific hypersensitive sites, the chromatin of active eukaryote genes is characterized by a generalized increase in sensitivity to nuclease digestion, thought to reflect its 'open' configuration. Villeponteau *et al.* reported last year¹² that when chicken red blood cells are treated with novobiocin the increased DNase I sensitivity of their β -globin genes is lost, which was interpreted as suggesting that torsional stress, induced by topoisomerase II, is required to maintain 'open' DNase-sensitive states of chromatin. As a counter to the criticism that novobiocin may inhibit enzymes other than topoisomerase II, Villeponteau reported in Cold Spring Harbour that γ -irradiation, which induces DNA strand cuts and can be presumed to relax torsionally stressed DNA, has an effect similar to that of novobiocin on the DNase sensitivity of active genes in chicken red blood cells. That topoisomerase II is in-



A mitotic chicken cell stained for DNA and for the 'scaffold' protein (Sc-1) that turned out to be topoisomerase II (with fluorescent labelled anti-SC-1 antibodies). The diffuse labelling represents radially projecting fibres of the cell's chromatin, and the brightly labelled regions show the presence of topoisomerase II at the bases of these loops (from ref. 18). Bar, 1 μ m.

involved in establishing 'open' states of chromatin is also suggested by A. Worcel (University of Rochester), who has shown that DNA either injected into *Xenopus* oocytes or in an oocyte cell-free extract can be assembled into torsionally stressed minichromosomes, but that these molecules are relaxed in the presence of novobiocin¹³⁻¹⁵. However, it must be emphasized that there is no direct evidence that eukaryotic topoisomerase II can induce torsional stress in DNA, and whether even active genes are torsionally stressed *in vivo* remains a matter for debate¹⁶.

A further line of evidence implicating topoisomerase II in eukaryotic gene expression has come from a study of RNA polymerase III transcription *in vitro*. J. M. Gottesfeld (Scripps Clinic) reported that the transcription of 5S RNA and tRNA genes in *Xenopus* cell-free extracts is inhibited by both novobiocin and anti-topoisomerase II antibodies. It seems that topoisomerase II is required for the assembly of stable RNA polymerase III transcription complexes which, once formed, are resistant to the effects of topoisomerase II inhibition. *In vivo*, however, inhibitor studies suggest that topoisomerase II activity is required continuously for the maintenance of RNA polymerase III transcription complexes.

Scaffolds

But perhaps the most startling discovery, which has come quite by chance from a very different line of enquiry by W. C. Earnshaw and his colleagues, is that topoisomerase II is among the proteins that make up chromosome 'scaffolds'¹⁷. These are the structures that remain after the most soluble proteins have been extracted from nuclease-digested mitotic chromosomes, and which are believed to reflect the organization of the chromatin into transcriptional units. By using fluorescently labelled antibodies, Earnshaw *et al.* have shown that topoisomerase II is located at the bases of the radial loop domains of mitotic chromosomes¹⁸ (see figure). If, as has been suggested, these loops represent topologically-constrained domains of DNA, then topoisomerase II would seem to be ideally

placed to regulate their topological state, and perhaps thereby their transcriptional activity. These results are consistent with the observation reported at the meeting by S. Elgin (Washington University, St. Louis) that antibodies to topoisomerase II uniformly stain *Drosophila* polytene chromosomes, and the finding of Berrios *et al.* that topoisomerase II is a major component of the *Drosophila* nuclear matrix¹⁹. Elgin's group had earlier reported that antibodies against *Drosophila* topoisomerase I show the enzyme to be concentrated at active regions of polytene chromosomes²⁰, suggesting that the two types of topoisomerases have quite distinct functions in *Drosophila*.

A number of lines of evidence, therefore, suggest that topoisomerases may be involved in regulating gene expression in 'higher' eukaryotes. These results are clearly in conflict with the unequivocal evidence that yeast genes continue to be transcribed in the absence of any topoisomerase II activity. Time will tell whether this reflects a fundamental difference between yeast and higher eukaryotes or simply a difference in experimental techniques. □

1. Fisher, L.M. *Nature, News and Views* **294**, 607 (1981).
2. Fisher, L.M. *Nature, News and Views* **307**, 686 (1984).
3. Goto, T. & Wang, J.C. *Cell* **36**, 1073 (1984).
4. DiNardo, S., Voelkel, K. & Sternglanz, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2616 (1984).
5. Uemura, T. & Yanagida, M. *EMBO J.* **3**, 1737 (1984).
6. Holm, C., Goto, T., Wang, J.C. & Botstein, D. *Cell* **41**, 553 (1985).
7. Trhasky, C., Voelkel, K., DiNardo, S. & Sternglanz, R. *J. biol. Chem.* **259**, 1375 (1984).
8. Han, S., Udvardy, A. & Schedl, P. *J. molec. Biol.* (in the press).
9. Udvardy, A., Schedl, P., Sander, M. & Hsieh, T. *Cell* **40**, 933 (1985).
10. Yang, L., Rowe, T.C., Nelson, E.M. & Liu, L.F. *Cell* **41**, 127 (1985).
11. Elgin, S. *Nature, News and Views* **309**, 213 (1984).
12. Villeponteau, B., Lundell, M. & Martinson, H. *Cell* **39**, 469 (1984).
13. Ryoji, M. & Worcel, A. *Cell* **37**, 21 (1984).
14. Glikin, G., Ruberti, I. & Worcel, A. *Cell* **37**, 33 (1984).
15. Ryoji, M. & Worcel, A. *Cell* **40**, 923 (1985).
16. Lilley, D.M.J. *Nature, News and Views* **305**, 276 (1983).
17. Earnshaw, W.C., Halligan, B., Cooke, C.A., Heck, M.M.S. & Liu, L.F. *J. Cell Biol.* **100**, 1706 (1985).
18. Earnshaw, W.C. & Heck, M.M.S. *J. Cell Biol.* **100**, 1716 (1985).
19. Berrios, M., Osheroff, N. & Fisher, P. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4142 (1985).
20. Fleischmann, W.C. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 6958 (1984).

Geoffrey North is Deputy Biology Editor of *Nature*.

Erratum

In N. E. White's article "A new breed of oscillators" (18 July, page 210), paragraph 3, line 13 should read $\sim 10^{12}$ Gauss and line 18, ~ 1 ms. In paragraph 7, line 2 should read ($\sim 10^{38}$ erg s⁻¹).

Children judge by the length

SIR—McGonigle in *News and Views*¹ refers to Piaget in support of the statement: "Asked to determine which of two rows of objects has 'more', four-year-old children tend to base their judgement on the shape of the set, selecting the more extended row and thus appearing to treat discontinuous quantities as continuous ones."

My recent research² questions whether the children are in fact treating the set as continuous. When three- and four-year-olds were provided with a number display while they counted the two rows, they correctly chose the shorter, more numerous row as having 'more'. The number display helps the children remember not only the number sequence but also how many objects were in each set.

Counting provides a less tangible cue to number than does a perceptual cue of length. When children select the longer row inappropriately, it may not be because they are treating the quantity as continuous: they may understand the discontinuous nature of the set but judge their own counting as a less reliable source of information about number than their judgement of length. Increasing the reliability of their counting by providing a visual aid increases the use of counting as a cue to judging number. Similarly, if children are provided with feedback showing them that counting is a more reliable cue to number than is length of row, they spontaneously start to count more often³.

SUSAN MICHIE

*Psychology Department,
The Royal Free Hospital,
London NW3, UK*

1. McGonigle, B. *Nature* **315**, 16-17 (1985).
2. Michie, S. *Br. J. educ. Psychol.* **54**, 245-283 (1984).
3. Michie, S. thesis, Univ. Oxford (1984).

Iridium, impacts and para-volcanism

SIR—If the iridium (Ir) anomaly reported¹ at about the Permian-Triassic boundary proves reproducible, it will provide another sort of evidence against the hypothesis that the Ir anomaly at the Cretaceous-Palaeogene boundary is due to an impact of an extraterrestrial body.

These two boundaries are at, or very close to, the culmination of what may be the two (or two of the three) largest regressions of the sea since at least the early Palaeozoic, each regression lasting several million years². If it is difficult³ to accept a coincidence in one case, the difficulty is squared for both cases together. With the reported finding by Kyte of no detectable anomalies after the Cretaceous in one core^{4,5}, although with a moderately high threshold of detectability, these two cases are the only ones known where Ir anomalies are presumably more than re-

gional. [The Permian-Triassic section sampled is in the only region known to have a more or less complete sequence, so the greater geographic extent of the anomaly must remain only a plausible hypothesis; nevertheless this anomaly is not known to be absent anywhere. The late Devonian Ir anomaly⁶ was concentrated by a blue-green alga and can be compared to the very much greater concentration reported⁷ in a kerogen in the early Late Permian Kupferschiefer. Another study of the Late Devonian elsewhere found no Ir anomaly (ref. 8, quoted in ref. 9).]

Obviously an impact could not have caused a regression which long preceded it in time, although it might somehow trigger an extended transgression. It is, though, conceivable that unusual paravolcanic activity could be coupled to a regression, although I know of no mechanism for either the coupling or (if one centre is involved) for the widespread distribution of ejecta.

I thank F.T. Kyte for comments.

LEIGH M. VAN VALEN

*Biology Department (Whitman),
The University of Chicago,
915 East 57th Street,
Chicago, Illinois 60637, USA*

1. Xu, D. *et al. Nature* **314**, 154-156 (1985).
2. Hallam, A. *Nature* **269**, 769-772 (1977).
3. Newell, N.D. *Geol. Soc. Amer. Spec. Pap.* **190**, 257-263 (1982).
4. Kyte, F.T. *Meteoritics* **19**, 257-258 (1984).
5. Kerr, R.A. *Science* **277**, 1453 (1985).
6. Playford, P.E., McLaren, D.J., Orth, C.J., Gilmore, J.S. & Goodfellow, W.D. *Science* **226**, 437-439 (1984).
7. Kucha, H. *TMPM Tschermaks Mineral. Petrogr. Mitteil.* **28**, 1-16 (1981).
8. Thierstein, H.R., *Geol. Soc. Amer. Spec. Pap.* **190**, 385-399 (1982).
9. McGhee, G.R. Jr, Gilmore, J.S., Orth C.J. & Olsen, E. *Nature* **308**, 629-631 (1984).

The fall of the Grand Menhir

SIR—Bahn¹ has alerted us to the recent discovery by Le Roux² that the capstone of the passage grave of La Table des Marchands at Locmariaquer in Brittany and a capstone of the dolmen at Gavrinis fit together and were once part of a single menhir which, when joined to another apparently related granite fragment used as a capstone in the long tumulus of 'er-Vinglé', north of Le Grand Menhir Brisé, would have stood about 14 m high, 3.7 m wide and 0.8 m thick and weighed about 100 tons. Carved decorations on the Gavrinis capstone include a *hache-charrue* (axe-plough, or hafted axe with a loop on the back of the haft), of which E. Shee Twohig states there are only five known examples³ including the one found by Minot⁴ on the Grand Menhir Brise, which lies broken in four fragments near Table des Marchands.

If we accept the authenticity and interpretation of the carving on the Grand Menhir Brisé as a *hache-charrue*, Le Roux's discovery requires reconsideration of previous suggestions as to the date and

function of the Grand Menhir, including the one made by A. S. Thom and me in 1980⁵ that the Grand Menhir, over 20 m long, was originally quarried, shaped, transported and erected to serve as a universal foresight, standing above the horizon when viewed from any direction, and may have been used by its erectors as a lunar observatory to predict eclipses. Taking into account changes in the obliquity of the ecliptic and the location of existing structures which could have been used as backsights, we suggested a possible date of erection of about 3700 BP. We concluded that the Grand Menhir probably fell either due to seismic activity or to a very powerful wind and lightning explosion at an unknown date prior to AD 1483.

It has been estimated that the dolmen at Gavrinis may have been constructed about 5200 to 5000 BP^{1,2}. Thus, the Gavrinis-Locmariaquer menhir probably was shaped and raised at least 5,000 years ago.

In view of Le Roux's new discovery, it is possible that the Grand Menhir was shaped and raised about the same time as the Gavrinis-Locmariaquer menhir or earlier, and may have been felled by the same force that felled the Gavrinis-Locmariaquer menhir. However, if the Grand Menhir fell at the time of the fall of the Gavrinis-Locmariaquer menhir, why were only fragments of the latter menhir reused as capstones in the construction of burial chambers, and the fragments of the Grand Menhir left undisturbed? Perhaps the answer is that the Grand Menhir fell sometime after the fall of the Gavrinis-Locmariaquer menhir and after the period of chamber tomb construction at Locmariaquer and its vicinity.

Is there a connection between the Gavrinis-Locmariaquer menhir and the Grand Menhir? It is possible that the erectors of the Gavrinis-Locmariaquer menhir had intended it to stand above the horizon when viewed from all directions and found it to not be high enough. They may have then erected the Grand Menhir sufficiently high for that purpose.

R. L. MERRITT

*2 Bratenahl Place,
Bratenahl,
Ohio 44108, USA*

1. Bahn, P.G. *Nature* **314**, 671 (1985).
2. Le Roux, C.-T. *Bull. Soc. préhist. Fr.* **81**, 240-245 (1984).
3. Shee Twohig, E. *The Megalithic Art of Western Europe*, 60 (Clarendon, Oxford, 1981).
4. Minot, R.S. *Bull. Soc. polymathique Morbihan* 89-98 (1965).
5. Merritt, R.L. & Thom, A.S. *Archaeol. J.* **137**, 27-39 (1980).

Powers of ten wrongly expressed—Corrigendum

IN this piece of Scientific Correspondence (by C. Liébecq, *Nature* **314**, 586; 1985) the last-but-one line should read: 'right-hand side equation (2) by 10⁶ (not 10³ as printed).

One man's sociobiology

T.H. Clutton-Brock

Social Evolution. By Robert Trivers.
Benjamin/Cummings: 1985. Pp.462. Pbk \$19.95, £19.95.

BETWEEN 1971 and 1976, Robert Trivers produced six seminal papers that established much of the theoretical framework of behavioural ecology and sociobiology. Trivers applied and extended evolutionary theory to explain the evolution of social and reproductive behaviour in animals, from arthropods to man. The importance of his work was quickly and widely recognized but the evangelical flavour of sociobiology, and its extension to human behaviour, led to conflict and controversy, especially in Trivers's home base at Harvard. In 1978, he left Harvard for the University of California at Santa Cruz and virtually ceased to publish, though it was widely rumoured that he was working on a new book.

Thus *Social Evolution* has been long-awaited. It is written as an introduction to the subject and assumes no previous biological knowledge. The first five chapters deal with the history of evolutionary theory, natural selection and group selection, basic genetics and kinship theory. The subsequent eleven cover kinship, reciprocal altruism and the evolution of cooperation; parental investment and parent-offspring conflict; the evolution of sexual reproduction, the sex ratio and sex differences in mortality; mate choice and deception. The book is designed to be readable. References are relegated to bibliographical notes, no formal mathematics is included and there are many illustrations which have been carefully selected to show animals performing the activities discussed in the text.

Social Evolution is very much one man's view of the state and prospects of sociobiology — but it is a brilliant view. Trivers combines great directness of thought with an urgent perception of the fundamental issues at stake. His book provides an articulate survey of the most important theoretical developments in the field which dodges no issues and pulls no punches. His writing consistently captures the excitement of recent discovery and the lure of unexplained phenomena, whether he is discussing the politics of alliances between chimpanzees or alternative solutions to the Prisoner's Dilemma.

Anyone hoping for another explosion of sociobiological theory was bound to be disappointed. The bulk of the book reviews evidence from established theories, concentrating on a limited number of studies which are described in detail, often supported by unusual examples drawn from a vast and diverse literature. However, there are few chapters that do not offer

some novel theoretical insight — for example, the extension of kinship theory to competition between genes, chromosomes and gametes; the consideration of parental interests in conflicts between siblings; the implications of female choice for male characteristics that benefit daughters but not sons; and the relevance of shifting selection pressures arising from coevolution to the evolution of sexual reproduction, dispersal and inbreeding. Two chapters also include important comparative reviews (of the distribution of paternal care and sexual reproduction) that have crucial implications for theoretical explanations.

As an introductory book, *Social Evolution* provides an enthusiast's guide to sociobiological theory illustrated by up-to-date examples and a stimulating tour of the most important controversies in the field. Many of its strengths lie in its idiosyncracies: its concentration on important problems rather than broad-brush coverage, Trivers's obvious delight in obscure examples and unusual applications of theory, and his readiness to provide an enthusiastically partisan account of academic controversies. But these characteristics have their disadvantages. Some topics of fundamental importance receive perfunctory coverage, among them the evolution of sociality itself, animal com-

munication and frequency-dependent models of conflict strategies. Devil's advocacy for functional and genetic explanations reaches extreme proportions in some cases. (Does Trivers really believe that human parents generally prefer their offspring to enter the celibate priesthood rather than engage in homosexual relationships because this increases the chance that they will invest in their kin rather than their mating partners, or is he joking? If so, it is a risky laugh.) Moreover the beautiful simplicity with which he illustrates theoretical arguments masks some messy facts — that male moorhens are 25 per cent heavier than females, despite female preferences for small males; that evidence indicating that vertebrates do not vary the sex ratio in relation to the population sex ratio, hatching order or paternal attractiveness is at least as strong as evidence that they do so; that a wide variety of field studies have found no indication of consistent mate choice by females. A substantial part of the book is not consensus science, but then it would have been disappointing if it had been.

I have no doubt that *Social Evolution* will be read by everyone with a serious interest in sociobiology or behavioural ecology and that it will be widely used as a course text. Trivers's greater achievement is that he has produced a book that will have an important influence on the development of theoretical issues in the field that he did so much to create ten years ago. □

T. H. Clutton-Brock is a Royal Society Research Fellow in the Large Animal Research Group, Department of Zoology, University of Cambridge, 34A Storeys Way, Cambridge CB3 0DT, UK.

Part of the problem

Ian Smart

Energy: Crisis or Opportunity? An Introduction to Energy Studies. By Diana Schumacher. Macmillan, London/Scholium International, 265 Great Neck Road, Great Neck, New York: 1985. Pp.335. Hbk £25, \$39.95; pbk £12, \$19.95.

DIANA Schumacher has designed a background reader for sixth formers, first-year undergraduates and non-specialists in energy studies. They will benefit from its stock of facts, and possibly from its emphasis on the ethical dimension of policy design. But they will be wise to treat it as serving a narrower purpose than it claims. Which is to say that Ms Schumacher's compendium needs both addition and admixture. She promises to lay out "the total context of the energy problem" (p.xiv). But her chosen audience should not suppose that this book alone provides a balanced diet.

The principal author has recruited expert help. Israel Berkovitch writes an admirable chapter on coal, Ron Hesketh a clear technical explanation of nuclear energy, and Judith Stammers an understandably committed brief on solar power. However, all the other sections come from Diana Schumacher herself — on energy history, oil and gas, biomass, geothermal and wind sources, water-borne energy, conservation and future policy options. Moreover, the main focus on Britain does not rule out frequent global expeditions. And there is some welcome attention to history throughout, despite obeisance to Amory Lovins's well-worn dictum about the future not being "the past writ large". Altogether, with such an agenda, it is no surprise that one person falls short of meeting the initial prospectus. In several minor respects and one major one, however, failure reduces the book's usefulness.

Most of the data are apposite and well-organized, partly as a result of choosing good sources. However, the stress is heavily on energy supply, especially from

"renewable" sources. With the exception of some introductory history and a rather conventional chapter on conservation, we are told little about the present structure or plausible elasticity of energy demand. And even the treatment of supply is irritatingly incomplete, in that we learn how much primary energy sources might notionally produce, but rarely how much they do in fact provide or could credibly supply at some specific future date.

Anyone recommending the book to students will want to balance it with some antidote for those ailments. But that is a slight matter in relation to its principal weakness: the absence of any basis on which a reader can compare the supply and policy alternatives discussed. In choosing between, say, different electricity fuels, a normal approach would be to set out their respective costs, including as necessary any "social" or "intergenerational" costs. No one reading Diana Schumacher's work will be prompted in that direction, simply because both broad and narrow questions of economic cost and benefit receive hardly more than an occasional — and commonly dismissive — allusion.

There is good evidence for the omission being a considered one. For example, "the energy future . . . transcends the realm of economics alone. . . decisions about it should be the democratic product of the public will" (p.293). Fair enough. In practice, however, Ms Schumacher imports other measures not to supplement economic analysis but to supplant it. Sadly, the effect is to undermine not only the value of her frequently informative book but also her occasionally moving case for changing the content of energy policy and the way it is made. Achieving beneficial change democratically depends on first equipping energy producers and consumers to judge what changes will, in fact, benefit them and their social and physical environments — which is *par excellence* a task for economically as well as socially sensitive educators. By leaving choices to be made only between what is old but familiar and what is new but uncertain, without offering any common measuring rod, Diana Schumacher resigns the energy future to "the devils we know". □

Ian Smart, an independent adviser on international energy affairs, 3 Grosvenor Avenue, Richmond, Surrey TW10 6PD, UK, is Vice-Chairman of the British Institute of Energy Economics.

Second editions

● *Rice Diseases* by S.H. Ou, published by Commonwealth Agricultural Bureaux, Farnham House, Farnham Royal, Slough SL2 3BN, UK. Price is £38 in CAB member countries, \$70 elsewhere.

● *Modern Practice of Gas Chromatography* edited by Robert L. Grob. Publisher is Wiley, price is £75.20, \$86.45.

● *Radiotelescopes* by W.N. Christiansen and J.A. Högbom, published by Cambridge University Press. Price is £30, \$59.50.

Understanding from Africa

W.T. Dean

Lower Palaeozoic of North-Western and West-Central Africa. Lower Palaeozoic Rocks of the World, Vol. 4. Edited by C. H. Holland. Wiley: 1985. Pp.512. £69, \$130.

PLINY the Elder's oft-quoted comment on Africa as a perpetual source of novelties assumed an especial significance for geologists everywhere during the 1950s and 1960s. Particularly exciting was the discovery, and the description by P. Hupé, of abundant, well-preserved Lower Cambrian fossils in the Anti-Atlas area of Morocco. Their importance lay in the fact that although the estimated age of the Earth in absolute terms is some 4,600 million years, it was only at a point about 570–590 million years ago (estimates vary) that the remains of animals with hard parts capable of being fossilized appear on a world-wide scale in the geological record.

In relative time, this point marks the end of the Precambrian and the beginning of the Phanerozoic, and in terms of rock successions coincides with the base of the Cambrian system, named for Wales. Unfortunately, in its eponymous country the Cambrian has not provided a readily definable base, but attempts to define this boundary have produced a remarkable amount of international collaboration and discussion, though no decision as yet. This digression will, I hope, serve to emphasize the importance of the Moroccan Lower Cambrian sections, though earlier hopes of them have not yet been fully realized.

It is clear from the preface of this volume that parts of the manuscript were delayed, and an unfortunate result is that some important, newer results are omitted. Among them are those of F. Debrenne from her work on the archaeocyathids, and in particular K. Szalay's preliminary report on trilobites still older than those described by Hupé. I may seem to be lingering unduly over Morocco, a country almost twice the size of Britain, at the expense of Algeria which is five times larger again. Nevertheless it is in Morocco, particularly the Anti-Atlas which provided the standard Lower Palaeozoic succession for the country, that documentation is most detailed.

It is inevitable that a book containing accounts of individual countries by different authors will lack coherence, but the contributions do have certain stratigraphic themes in common and a unifying summary or review would have made the volume more generally useful. One of the benefits of plate tectonic theory is that it has rendered respectable the sort of palaeogeographic reconstructions, includ-

ing those for the Lower Palaeozoic, that would have been unacceptable 30 years ago. At a time when Africa formed part of an even larger continental mass, marine platform sedimentation took place on a scale unparalleled in modern times and relatively modest sea-level changes had far-reaching consequences in terms of vast marine transgressions and regressions. The striking absence from Africa of Upper Cambrian rocks is part of a wider phenomenon involving regression, not always complete, in the Middle East and parts of Europe, and some discussion of it would have provided additional interest.

For many geologists the demonstration in 1963 that a large glacial complex some 440 million years old, centred on the late Ordovician South Pole, lay in what is now the desert areas of southern Morocco, proved at least as exciting as the new data for the Lower Cambrian. It would have been helpful for general readers to be told something of the accompanying regression, whose effects are claimed to have influenced both sedimentation and faunas as far afield as southern Scandinavia and parts of present-day North America, then much closer to Africa. Marine transgression on a grand scale followed the end of the glaciation, though the age of the resulting basal deposits may vary slightly with location. Again, the regional story involves not only Africa but also parts of the Middle East and Europe, and there are highly relevant researches, such as those of J. K. Leggett on associated deposition of black shales, that one would have liked to see mentioned.

These points seem minor when set against the overall benefits of this first general review, in English, of a region of daunting size and often difficult terrain. Many readers may not appreciate that even in Britain, which has one of the world's oldest national geological surveys, systematic geological mapping of all parts of the country has still not been undertaken. But the size of Morocco and Algeria is far exceeded by that of west and west-central Africa, a region less well known and the subject of an admirable review that contains much previously unpublished information, including that from oil exploration.

The book's contents are in large measure a tribute to the efforts over many years of numerous French and, in west-central Africa, Belgian geologists, whose work is evident in the full reference lists. The volume is dedicated, appropriately, to the late Henri Holland, doyen of workers in Morocco. It is particularly sad to read on the book's dust-cover that this is the fourth and last of a series that was still far from completion and for which there remains a widespread need. □

W. T. Dean is a Professor in the Department of Geology, University College, Cardiff, PO Box 78, Cardiff CF1 1XL, UK.

The past in current application

W.D. Hackmann

Electricity and Medicine: History of Their Interaction. By Margaret Rowbottom and Charles Susskind. *Macmillan, London/San Francisco Press, Box 6800, San Francisco, California: 1984. Pp. 303. £25. Price in the United States not known.*

IN 1743, so we are told by a contemporary German source, Johann Gottlob Krüger, who was professor of philosophy and medicine at the Lutheran University of Helmstädt, was asked by a student what possible use there could be for electricity — then the new rage in all the drawing rooms of Europe. After a moment's consideration, he replied that as it could be of no possible value in theology and jurisprudence, there remained nothing left but medicine. The first "practical" application, indeed, was in medicine, long before the discovery of the incandescent lamp or the electric motor; and thus began the long association between electricity and medicine, chronicled in detail in this richly illustrated book.

This fascinating story has been largely ignored until recently and the reasons are not hard to fathom. The early history of electrotherapy, especially, contains elements that appear distinctly unscientific, certainly do not fit easily into the conventional view of what science is about, and the treatments have (with hindsight) little modern theoretical significance. Those who were primarily concerned with the pedagogic value of the history of medicine saw little point in burdening their students, or themselves, with the topic.

Besides electrotherapy, or "medico-electricity" to use an eighteenth-century term, Rowbottom and Susskind have examined many other aspects of the interaction between electricity and medicine. Six broad periods can be identified in their book. The first is the use of frictional electricity, or electrostatics, from the 1740s until the end of the eighteenth century. This was followed by the application of current electricity or "galvanism", after Volta discovered the electrochemical battery in 1800, which in turn was superseded by the induced current or "faradization" of Golding Bird, Duchenne and others. Up to this time, electricity was seen as a cure-all for a vast range of complaints, ranging from consumption to venereal disease and from nervous headaches to rheumatism. Treatments were based on contemporary physics, physiology and medical theories, and on a corpus of experimental work on the reactions of organisms to electricity. An important development of the mid-nineteenth-century was the rise of electrophysiology. The application of electricity, too, became more

specific, thanks in the first instance to Duchenne and Remak. It began to be used increasingly in clinical medicine and physiotherapy (in the treatment of damaged muscle tissue), as an aid to diagnosis with a large variety of electrical devices, and in electrosurgery. The final three broad periods, which bring this history up to the present, are the application of high-frequency currents (begun by D'Arsonval in the late 1880s) and of electromagnetic radiation such as X-rays and radioactivity, and the impact of electronics and the computer on twentieth-century medicine.

The book is largely a factual and very detailed account of these multifarious developments. The narrative is rather impersonal and bland, in the style of a lengthy report, but even so the fascination of the subject comes through. What is still



Electrical treatment in the late eighteenth century — passing a discharge through a part of the patient's body. (From W. van Barneveld's *Geneeskundige Electriciteit*, published in Amsterdam in three volumes between 1785 and 1789.)

required, however, is a deeper historiographical analysis. Matters that emerge from the book, and which could be explored further, are the speed with which new discoveries in physics found their way into medicine, even when the scientific principles involved were not fully understood (a good example is ionic medication), the significance of the underlying philosophical and scientific frameworks in the formulation of the medical problems and in the acceptance of new forms of treatment, and the importance of the placebo effect. For instance, eighteenth-century electrotherapy and Mesmer's "animal magnetism" had a common ancestry; the former is now an important tool in neurophysiology, while the latter has evolved through Charcot and Freud as a means of studying the human psyche. These aspects, hinted at in Rowbottom and Susskind's pioneering work, are well worth further study. □

W. D. Hackmann is Assistant Curator of The Museum of the History of Science, University of Oxford, Old Ashmolean Building, Broad Street, Oxford OX1 3AZ, UK.

Stranger to physics

Thomas F. Glick

The Young Einstein: The Advent of Relativity. By Lewis Pyenson. *Adam Hilger: 1985. Pp. 246. £19.95, \$28.*

LITTLE more can be learned about the cognitive structure of the origins and development of special relativity by Einstein and his early followers. It is fruitful, however, to explore the cultural and social milieu in which the young Einstein lived and worked, in search of a greater understanding of the context in which the theory of special relativity was formulated. In this effort lie both the strengths and limitations of Pyenson's approach—strength particularly in his examination of the multifaceted relationship between mathematics and physics in Wilhelminian Germany, limitation because the focus upon Einstein tends to become blurred in such a broad-gauged analysis and because it is difficult to pin down the precise intellectual roots of Einstein's early development as a scientist, except by inference.

The first three of the nine essays comprising the volume deal with specifically biographical materials: how Einstein's world-view was influenced by his education, his family's electrical machinery business and his Jewish roots, respectively. The identification of Einstein as a loner—*Einspänner*—in the context of German Jewry is wholly believable (while hardly surprising), as is the application of Simmel's notion of the social "stranger" and Park's "marginal man". Less credible is the possible effect on Einstein's mentality of the family electrotechnical business; such surroundings may have whetted young Albert's appetite for physical questions, yet Pyenson's speculation that Jakob Einstein's clock-meters were a possible stimulus in the genesis of special relativity seems to be stretching the evidence to its maximum. The case for the influence of secondary-school education on the mature scientist, although inferential, is nevertheless quite strong, for it rests on understanding the nature of the changing relationship between mathematics and physics in the 1890s, when Einstein was a student.

Most compelling is Pyenson's discussion of the notion of the pre-established harmony between mathematics and physics as an ideology which informed much of the discussion in special relativity. Leibniz's doctrine, stripped of its theological underpinnings, became, in the hands of pure mathematicians, both a justification for pure mathematics and a support for acausality in physics. By applying the concept, mathematicians, in Pyenson's view, were able to "de-relativize" the special theory and cast it as a theory of absolute space-time. Against this background,

Pyenson understands the approaches to relativity of a whole succession of commentators from Einstein's early collaborator Jakob Laub, to Born and Weyl, the latter an extreme exponent.

In a discussion of Minkowski and the special theory, Pyenson notes that, as a student, Einstein had found him boring as a teacher, because of the lack of physical content in his lectures, and that, later, Minkowski judged Einstein as mathematically limited. A recurrent theme here is

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Genius matured — Einstein, aged 37, in his Berlin study in 1916. The photograph is reproduced from one of the introductory chapters in Jamie Sayen's Einstein in America: The Scientist's Conscience in the Age of Hitler and Hiroshima. Publisher is Crown, New York, price is \$17.95.

that the tension between pure mathematics and theoretical physics was a constant feature of the scientific atmosphere in which relativity emerged. Similarly, the members of the famous 1905 Göttingen seminar on the electron theory were unable to perceive what was new in special relativity, and to resolve the impasse in the electron theory which interested them for its intrinsic and abstract characteristics rather than as a problem in physics. Einstein was opposed to this approach to physics which began with pure mathematical constructions rather than with physical problems.

Pyenson takes the question one step further by examining the relationship of mathematics and physics in secondary-school education in order to explain how pressure "from below" may shape the attitudes of research scientists. In this view,

students in traditional *Gymnasien* learned mathematics from people who were the products of university seminars in pure mathematics and who stifled generations of students with abstractions. When engineers and others revolted against this approach in the 1890s, a new instrumentalist viewpoint emerged, especially in the *Realenstalten* which prepared students for practical careers in technology and industry and which also, after 1900, produced the majority of physicists. The new instrumentalism, epitomized by Felix Klein, was a middle position between the older view of mathematics as a mere tool and the emergent neo-Leibnizian current espousing pre-established harmony. Einstein's education at the Cantonal school of Aarau was closer in spirit and content to the programme of the *Realenstalten* than to the classical *Gymnasium* curriculum which he had found so restricting.

It is not quite clear in Pyenson's exposition just how the pedagogical revolution was achieved. The secondary-school mathematics teachers (here characterized as the *sans-culottes* of this revolution)

were themselves the products of training in pure mathematics. The engineers wanted to break the hold of the universities (and thus of the pure mathematicians) on the training of school teachers and backed educational reforms designed to bring this about. Klein's role was to co-opt the *sans-culottes* by bringing the instrumentalist, problem-solving approach into the university. The complex social dynamics at play here require further delineation.

In his recent book *Revolution in Science*, I. B. Cohen stresses the significance of the conversion of "revolutions on paper" to true scientific revolutions; the social processes whereby working scientists are converted to a new theory are crucial to the analysis of this transition. Pyenson's contribution is precisely the exploration of factors in the German educational system which pre-conditioned a range of responses to relativity. □

Thomas F. Glick is Chairman and Professor in the Department of History, Boston University, Boston, MA 02215, USA.

Marking change

Arthur J. Boucot

Fossils and Climate. Edited by P. J. Brenchley. Wiley: 1984. Pp. 352. £27, \$49.95.

I ENJOYED reading this book, and wish I could have attended the symposium on which it is based. As is usual with such volumes, the 15 contributions differ in both quality and originality, but in general the standard is very high.

For me, the outstanding chapters are Janis's on the use of fossil ungulate communities in the Cenozoic to better understand floral communities of the past, as well as climate, and Rögl and Steininger's synthesis of the Paratethyan story ("Pontian Cockles" brought up to date). This latter contribution is clearly the culmination of some decades of labour by scientists from many nations, and will be of great value to both geologists and palaeontologists, as well as evolutionists. Not far behind is Rosen's account of just what might control reef coral distributions, both past and present. Rosen discusses our present understanding of the subject, much of which is based on his own research, and evaluates critically the various explanations which have been put forward (one hopes that he will next tackle the Cenozoic record to complement this work). Also included is a helpful commentary on the thorny problem of defining biogeographical provinces.

In addition to these particularly original contributions, there are several less innovative but still very worthwhile chapters. In this category are Shackleton's

account of the use of oxygen isotope evidence in the study of Cenozoic climatic change, Creber and Chaloner on growth rings in fossil wood, Hallam on the use of marine invertebrates for interpreting past climates, Brenchley on the relation between the terminal Ordovician extinction event and climatic events (contemporary glaciation), and Dickins on the story of Gondwana in the Late Palaeozoic.

In the palaeobotanical area, Boulter reviews the drawbacks of using European Palaeogene botanical evidence in estimating temperatures; these are clearly important problems, of which I had been ignorant. Boulter doesn't provide us with a pleasing answer, however, possibly because one isn't available. Here, too, Batten reviews the confusing (for me) status of late Cretaceous pollen biogeography and its implications.

The book, of course, hardly lives up to its title, but it would be unreasonable to have expected this from a handful of papers prepared for an international symposium. Especially notable is the absence of any discussion of meteorological modelling of the past (CLIMAP, as well as work on the Cretaceous such as that by Barron or Lloyd), but the bibliographies at the end of each of the papers are reasonably comprehensive and do provide a lead into the broader literature on fossils and climate. Libraries specializing in geology and biogeography should purchase this book, while life scientists would do well at least to skim the papers which deal with the aspects of evolution involving climatic change and biogeography. □

Arthur J. Boucot is Professor of Geology and Zoology at Oregon State University, Corvallis, Oregon 97331, USA.

Innovation and academic research

from Donald W. Braben

Present policies on the commissioning of fundamental research seem likely to inhibit major innovation in the long term. An alternative approach may avoid this pitfall.

FEW would doubt that for Britain to survive as a major industrial nation, its industry must be transformed from the old labour-intensive technologies of yesterday to new idea-intensive technologies. The transformation will depend in part upon social, political, financial and commercial factors, but unless the technological options can be created, there will be no grist for our national mills, however they may be transformed.

If we accept the assumption that the flow of scientific discoveries is not about to dry up, how best can we tap it in a controlled way? For development—that is, the steady improvement and application of an idea—the answer is well known and much industrial research is indeed dedicated to this purpose. But is it possible to control access to the sort of major discoveries that have laid the foundations of today's high technology, such as those which led to the new electronics, the laser and genetic manipulation? It remains impossible, as it always has been, to predict specific revolutions. But since most scientists are confident, in general terms, that the future holds major surprises in store, the best strategy might seem to be to ensure that research is sufficiently diverse so that all future bets are covered, and to allow serendipity to take its course.

A strategy for diversity is attractive, as it implies that we need to take little action but to await events. Moreover, since most societies would naturally opt for such a passive approach, one might expect that it would be universally accepted. However, I would argue that although this strategy was in favour, at least implicitly, not long ago, the practice of ensuring that research has a rich diversity is now too rare, is still diminishing and ought to be increased.

Nowadays the traditional funding organizations are increasingly allowing their perceptions of national interest to influence their selection of research proposals. For short-term research and development, few would quarrel with this approach, but for fundamental, curiosity-driven research, which may lead to the major industry of the next century, I believe it may be counterproductive.

Innovation

A better approach to the problem of selecting the fundamental research that may precipitate major innovation would

be to acknowledge fully our inability to predict the future, to abandon the approach of supporting only those fields of research that are designated as priorities and to support research in any field provided certain criteria are satisfied.

The choice of the criteria for selection should come from an analysis of the methodology and philosophy of science rather than from the conventional approaches based on disciplines and specific problems. An important criterion in this respect is concerned with the way research is formulated. Long ago, there existed a subject called natural philosophy or natural science, but as our understanding has developed, most organizations involved with science have found it increasingly necessary to partition scientific endeavour into subjects of manageable size. Not so long ago, such simple divisions as physics, chemistry and biology were sufficient, but now the number of partitions is enormous. It now seems important to distinguish between chemical physics and physical chemistry; there are biophysical chemists, molecular biophysicists and physicochemical hydrodynamicists.

What is more, the practitioners in these fields usually take great care to define the breadth and scope of their discipline because often there is a journal publishing the results of research fitting that description, or a special committee to sponsor it within a national funding organization. From time to time, however, national funding organizations become aware that these boundaries are too finely drawn and new "interdisciplinary" subjects are defined and declared national priority areas to tempt researchers away from their old haunts. Thus in the past few years, we have seen in the United Kingdom the creation of such new fields as polymer science, marine technology, biotechnology and information technology. But however they may be defined, none of these groupings have tangible significance in the natural sciences.

Compartmentalization

With the possible exception of the so-called "big sciences", like particle physics or astronomy, the pressures to ensure that proposals for research fit neatly into the structure of the day are substantial and growing. But if we wish to stimulate the

research that may lead to the unexpected, we should be prepared to eliminate any artificial constraints on creativity and recognize that no matter how we draw boundaries, they will inhibit the creativity we are trying to foster. Thus, scientists should be free to formulate their research so that they may launch a coherent attack on the problem they have identified. Interdisciplinary research does not necessarily help, since that approach tends to paper over the cracks. It may not be a good idea for a physicist to collaborate with a zoologist. The present disciplines are an indication of our present understanding, and by implication, our prejudices. As these are precisely the things that researchers should be striving to change, it seems sensible to avoid any artificial constraints from the outset.

It is not, however, sufficient merely to have an expansive idea. Grandiose schemes are not at a premium. There should be credibility, and some prospects, however distant, of eventual innovation. What is most important is to identify research of the type that may lead to innovation, and to back those individuals or small groups who need freedom to develop their ideas fully.

How should we identify those people most likely to succeed? Of course, they should be highly motivated and creative people, who have original ideas of their own; but what else? The approach we are taking in the BP Venture Research Unit is in effect to invite researchers to respond to what seems to be a simple question — if you had all the resources you would need, what would you do that you are not doing now? In our turn, we look for proposals that focus on a substantial phenomenon or observation which is not understood, and where our lack of understanding is a serious obstacle to progress. We also look for and encourage novel and imaginative approaches aimed at increasing our understanding. The research should not be concerned with development, since we are looking for the unexpected and do not know what we want developed, and so should not have narrowly specified goals. On the other hand, the scientists should have specific ideas of the initial directions their research will take and be prepared to modify them as their understanding develops.

As we are searching for the unexpected,

we feel that the best we can do is to select that research which offers divergence or a rich spectrum of possibilities, at least one of which may have industrial relevance. But it is also important that scientists should be free to develop their ideas as they think fit without external direction.

Novelty or freshness of approach can come from a variety of sources. It can, of course, come from a new observation, or a new collaboration with someone from a different disciplinary background. It can also come from a change of job — a new chair or a new postdoctoral fellowship. Young people should be encouraged, particularly those who, having won their doctoral spurs, want to ride off over the horizon. However, they are usually discouraged in their attempts to gain substantial support for their own ideas, which is often a pity, because they are in a phase of their careers when their creativity is high and their respect for tradition has not yet matured — a combination of qualities which would seem ideal for fomenting revolution!

Coherence

We should also try to encourage those researchers who have reached a point of departure in their careers and wish to try something new, particularly if their approach is unique and "coherent". Coherence in this context refers to the way the research is formulated, and is used as a shorthand to describe conceptually formulated research where the central questions are addressed effectively and are logically well founded. How many of us have worked hard to answer specific questions only to find when we have the answers that we asked the wrong question in the first place? Negatively formed questions can be particularly hazardous. The question, for example, of why plants do not fix nitrogen directly may never be answered completely. On the other hand, it is well known that some bacteria know how to accomplish this and a better question might be — how can we achieve genetic transfer between bacteria and plants? Similarly, research based on a single hypothesis of the "what if ...?" type can also be very risky, since the hypothesis may fail because of some factor that has not been considered, and little will have been learnt. Research driven by a technique also seems to be a poor bet, since almost invariably the technician's skill is a solution looking for a problem. That may be fine for development, but in the search for the industries of tomorrow it is more important to look for perspectives that place the problem in hand in its proper context with respect to our understanding of the field in general, and which suggest to the scientists a range of techniques and hypotheses which may be tried initially. We use the term "coherent research" in this context, therefore, to mean that the scientists appear to have identified all factors that can contribute to their problem

and are therefore sufficiently armed to make a successful attack.

Synthesis

These remarks refer implicitly to the natural sciences. What of the sciences with a completely different character and structure, which I like to call the synthetic sciences? In many respects they have better names individually, such as engineering and design, but the images that those words usually conjure up in many minds differ radically from the activity we envisage, although in principle there is no reason why this should be so. From our point of view, engineering or design seem to be concerned with the marshalling of resources — ideas, people and equipment — for the achievement of specific goals. How should the resources best be put together to meet the requirements? The engineer may spend hours, weeks or even years thinking about this question and at some point will put pen to paper and begin to sketch out his final design. How does the engineer structure his thoughts during this time? What disciplines are brought to bear? They are certainly abstract and many would say that they boil down to experience and judgement; as is the case of the natural scientist contemplating a problem. The difference is that for the engineer, the only complexity eventually remaining concerns the interrelationships between the large number of factors contributing to the problem, and in this respect the complexity is entirely man-made.

The problem of how to control this type of complexity has been with us for a long time, but in the past few decades a number of formal methodologies have been derived which may eventually simplify the solution and provide a framework in which we will be able to produce designs more efficiently. These advances have come within the field of computing science, which in many ways may be seen as a microcosm of engineering in general. Computers can now be used to test formal approaches to the control of man-made complexity, and they allow us to experiment with new ways of thinking. As may be expected, the approaches are abstract and often involve new uses of old mathematics such as logic, algebra and calculi, which now offer the prospect of much higher reliability and efficiency in the production of computer programs, and in the longer term the prospect of more efficient engineering processes in general.

Engineers and designers are no strangers to mathematics, but they have tended to restrict their interests to analytical techniques and in general, have tended not to realize the potential of the mathematics that will enable them to manipulate concepts and ideas. Computer scientists are now beginning to make progress in this direction. The potential benefits of increased reliability and efficiency in computing are becoming widely recognized to-

day — the Royal Society recently held a conference on the importance of mathematical logic to computing. However, the wider implications of rigorous methodologies for engineering and design have not yet commanded the attention in the universities that the subject deserves, although the US National Science Foundation is about to launch a programme on the science of design. For our part, the Venture Research Unit will continue to search for people with novel ideas in what may also be called theoretical engineering.

To realize the full potential of an initial laboratory discovery will often require as much ingenuity as went into the creation of the opportunity; penicillin is a classic example. In anticipation of success, therefore, it is essential not to let the researchers participating in any scheme get lost to sight, so to speak. In our Venture Research schemes, we therefore discuss progress regularly with each team on a conceptual level and arrange occasional visits for the discussion of possible prospects with industrial experts.

The problems of developing fundamental discoveries for industrial application are well known, and they become acute when the ideas are novel and not extensions of existing interests. The key factor is to foster the transfer at every level from the bench to the market-place, and to try to ensure that everyone involved has something to gain from a successful transfer. That classic cliché for failure, "not invented here", can usually be interpreted to mean that there is nothing in it for the developer of that invention. Everyone involved in a development programme should have a real stake and stand to gain if the transfer succeeds.

Constraints

The universities in Britain and elsewhere today are suffering increasingly from financial constraints, and even distinguished researchers sometimes have difficulty obtaining adequate support for their fundamental research. The current tendency for national funding organizations to favour research based on extrapolations from present trends will inevitably have the effect of reducing diversity in national research programmes and possibly the rate at which breakthrough occurs. In my own experience, priority areas look depressingly similar and they all seem to be aiming for the same goals. That is why we aim to extend our support of fundamental, curiosity-driven research through the backing of master tacticians who also have strategic vision, and perhaps have the best of both worlds.

The comments on the manuscript from my colleagues Peter Beadle and David Ray are gratefully acknowledged. □

Donald Braben is Head of the Venture Research Unit at BP International Limited, Britannic House, Moor Lane, London EC2Y 9BU, UK

Can spores survive in interstellar space?

Peter Weber & J. Mayo Greenberg

Laboratory of Astrophysics, University of Leiden, 2313 AV Leiden, The Netherlands

Inactivation of spores (Bacillus subtilis) has been investigated for the first time in the laboratory by vacuum ultraviolet radiation in simulated interstellar conditions. Remarkably, damage produced at the normal interstellar particle temperature of 10 K is less than at higher temperatures, the major damage being produced by radiation in the 2,000–3,000 Å range. Our results place constraints on the panspermia hypothesis.

THERE is good evidence that life existed on Earth as early as 3,800 Myr ago¹, which leaves only ~200–400 Myr after the Earth's crust cooled² for terrestrial chemical evolution to lead to primitive, yet fully developed, life forms. That this might be too short a timescale for such an evolution has led to speculation that either cosmochemistry^{3,4} or even panspermia^{5,6} is involved in the origin of life on Earth.

We see no reason to reject the hypothesis that several hundred million years is adequate to produce the required prebiotic chemical precursors on Earth. No answer can be given to this question until the basic mechanisms of this transformation are understood. If the probability were quite small that life could start in 300 Myr, does extending this time by a factor of only 50, which takes us to the probable beginning of the Universe (some 15,000 Myr ago⁷), increase the probability sufficiently? Inverting this argument, we deduce that if life is certain to have evolved in 15,000 Myr, the probability cannot be negligible in 300 Myr.

We do not seek to address such questions here; they have been discussed elsewhere^{8,9}. If we accept the hypothesis that life began from primordial prebiotic molecules, whether formed first on the Earth² or coming from outer space⁴, does this belief exclude the possibility that life, once formed, can be transported from one source to another?

We reconsider below some of the ideas of Arrhenius¹⁰ by placing them in a modern astrophysical context. The four fundamental phases of panspermia are: (1) removal to space of biological material which has survived being lifted from the surface to high altitudes; (2) transport from one solar system to another; (3) survival of the biological material over timescales comparable with the interstellar passage time; (4) deposition of the biological material onto a new host planet's atmosphere and surface. We argue that phases (2) and (3) are amenable to quantitative treatment and defer consideration of phases (1) and (4), with phase (1) presenting more problems than phase (4).

Arrhenius¹⁰ suggested that solar radiation pressure could provide a mechanism for driving microorganisms into interstellar space with enough speed to reach another star in a 'reasonable' time which he considered to be 3,000 yr. He was not aware that microorganisms are killed almost as soon as they are exposed to the full solar ultraviolet (UV) on leaving the Earth's atmosphere. Furthermore, he ignored the reverse effect of the radiation pressure which acts to prevent a small particle from entering the environs of another star.

Insofar interstellar transport is concerned, there is an obvious mechanism through the random motion of molecular clouds. The gaseous and particulate matter between the stars is unevenly distributed into concentrations or clouds, with densities up to 10^4 hydrogen atoms cm^{-3} ($n_{\text{H}} = 10^4 \text{ cm}^{-3}$) and higher. These clouds move about at random speeds of 10 km s^{-1} (ref. 11). Thus, should a bacterial spore be captured into such a cloud it will be swept along with the gas. Given the distance between neighbouring stars of ~0.1–1 pc (0.3–3 light yr)¹², this corresponds to a passage time of 10^5 – 10^6 yr. Thus, if one star in 1,000 possesses a solar system, a survival time of 10^6 – 10^7 yr is required.

The question we address is can a spore survive as long as 1–10 Myr in a molecular cloud?

A major destructive mechanism for spores in interstellar space is the photolysis caused by photons from starlight. The flux of UV photons with energy $\geq 4 \text{ eV}$ is such that each molecular bond in a 0.5- μm radius particle will have been subjected to such a photon every 1,000 yr in the less dense regions of the space between the stars. In the densest regions, this timescale may be increased to 10^6 – 10^7 yr. Because the small particles in space are normally at temperatures of 10–15 K (refs 13, 14) at first it seems difficult to imagine the maintenance of the chemical integrity required to preserve a living organism.

The three basic factors in interstellar space which are hostile to microbes are: (1) vacuum. In interstellar space the densities are much lower than on Earth. Even so-called dense interstellar clouds with 10^6 hydrogen atoms cm^{-3} ($n_{\text{H}} = 10^6 \text{ cm}^{-3}$) have a pressure of only 5×10^{-13} mbar, which would be considered a very good vacuum. (2) Energetic photons and cosmic rays. Even though the average point in space is very distant from the nearest star, the UV radiation is still a factor. Although individual cosmic-ray protons, X rays or γ rays may be more lethal than UV photons, their total energy input is generally orders of magnitude lower than that by the UV. (3) Temperature. Although the temperature of the gas in interstellar space is generally in the range 10–100 K, the mean temperature of a small particle, for example, a bacterial spore or a virus, is generally $\leq 10 \text{ K}$. Thus, we have to consider what happens to a complex biological system in conditions of high-vacuum, low-temperature and vacuum UV irradiation.

Experimental method

We performed the work described here at a laboratory facility routinely used in the Leiden Astrophysics Laboratory to simulate interstellar conditions for the study of chemical evolution of interstellar grains by photoprocessing simple molecules^{15,16}.

There is both a similarity and an essential difference between interstellar grain evolution and interstellar effects on living organisms. In the former, we are interested in the formation of complex organic molecules from simple molecules, whereas in the latter, we are interested in the destruction and rearrangement of already existing complex molecules. In either case, the basic process starts with the breaking of a molecular bond or the ionization or excitation by a UV photon. The experiment is illustrated in Fig. 1. To relate our new data to earlier investigations, we selected organisms that have been studied previously in various conditions of irradiation, temperature and vacuum, but never in conditions that simulate the interstellar medium.

Bacillus subtilis spores are suitable for viability tests in simulated space environments relevant to the Solar System¹⁷. They are resistant to vacuum exposure; even in ultra-high vacuum ($\leq 10^{-10}$ torr) ~90% of spores maintain their colony-forming ability compared with desiccated spores. By means of high-vacuum chambers it is possible, using simulated interstellar conditions, to study which cell components are affected either by vacuum or photolysis, or both.

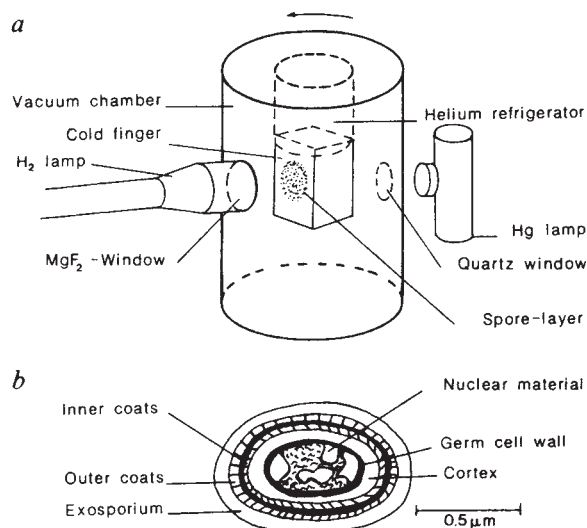


Fig. 1 a, Apparatus used to irradiate spores. The cold finger on which the spores are deposited is maintained at either 10 K or room temperature. The quartz window is also used in conjunction with the H_2 lamp. b, Cross-section of the spore; scale bar, 0.5 μm .

We limited our experiments to two strains of the *Bacillus subtilis* spore, one which is relatively irradiation-resistant, wild-type 168 (Nester, Marburg), and the other, TKJ 6323 (provided by N. Munakata), which is sensitized to irradiation as a result of repair deficiencies. The spores used in our experiments are freed of all vegetative cells and other debris by enzymatic treatment and density gradient sedimentation.

Results

In the first experiment, we investigated the variability of susceptibility of spores of different genotype to UV (wavelength 2,537 Å, from a low-pressure mercury lamp). This wavelength is termed RUV (reference UV). The parameters of our experiments are summarized in Tables 1 and 2. The inactivation kinetics for the sensitive and wild-type strains at room temperature and atmospheric pressure are compared in Fig. 2a to show the strongly enhanced sensitivity of the repair-deficient strain TKJ 6323 to RUV. Because of mutations in the excision-, recombination- and sporephotoproduct repair, strain 6323 is ~50 times more sensitive to RUV than wild-type strain 168 in aqueous suspensions and ≥ 10 times more sensitive when irradiated on a surface.

The effect of low temperature (10 K) and high vacuum ($<10^{-6}$ torr) on the amount of lethal damage to spores is compared with results at room temperature, whether at atmospheric pressure or high vacuum, in Fig. 2b, c. We find that strain 168 irradiated at room temperature is at least four times more sensitive in vacuum than at 1 bar. This is the same effect as found by Horneck¹⁷. If, however, the sample is cooled to 10 K before irradiation, spores are not inactivated with fluences which would have been lethal dose rates at room temperature.

Although an increase in sensitivity is often observed with decreasing temperature to $\sim -80^\circ\text{C}$, decrease in sensitivity to RUV is observed below 130 K (ref. 18). Extrapolating the data of Ashwood-Smith *et al.* we predict that a dose rate of $2 \times 10^4 \text{ J m}^{-2}$ is sufficient to inactivate samples to F_{10} , where F_{10} is defined as the fluence required to inactivate 90% of the spores. We confirm this trend to less sensitivity at extremely low temperature by noting that irradiating a sample of wild-type (WT 168) at 10 K and vacuum with a dose of $1 \times 10^4 \text{ J m}^{-2}$ of RUV allowed 25% survival.

An experiment performed with TKJ 6323 (Fig. 2c) indicates that this RUV-susceptible spore is as strongly resistant as WT 168. The amount of lethal lesions is significantly reduced at 10 K because TKJ 6323 is unable to repair damages during germination and development to the vegetative cell.

To simulate the UV flux in interstellar space, we use a micro-

Table 1 Comparison between laboratory and interstellar conditions

	Laboratory	ISM
Pressure	8×10^{-8} mbar	3×10^{-15} mbar
Temperature	≥ 10 K	≥ 10 K
UV flux ($E > 6$ eV)	10^{15} quanta $\text{cm}^{-2} \text{s}^{-1}$	10^8 quanta $\text{cm}^{-2} \text{s}^{-1}$
Fluence ratio:		
Far UV (2,000–3,000 Å)		
Vacuum UV (1,000–2,000 Å)	5	1.1

wave-powered H_2 -discharge lamp. The emission of this source is peaked, in the vacuum UV ($1,000 < \lambda \leq 1,900$ Å), at 1,215 Å (Lyman α) and 1,600 Å and has an increasing continuum in the far UV (2,000–3,000 Å).

The mean flux of the lamp's spectrum, in the vacuum UV, is of the order of 1.5×10^{15} quanta $\text{cm}^{-2} \text{s}^{-1}$. The energy flux of the lamp above 2,000 Å is roughly five times as high as in the vacuum UV portion. We thus have to use two different scales when demonstrating the vacuum UV irradiation and results with full spectrum irradiations (see Tables 1 and 2). Estimating a mean flux of vacuum UV photons in the average interstellar medium of $1 \times 10^8 \text{ cm}^{-2} \text{s}^{-1}$, 1 h of irradiation by vacuum UV in the laboratory corresponds to $\sim 10^3$ yr in space. The first question then is whether vacuum UV with photon energies > 6 eV ($\lambda < 2,000$ Å) has an effect on the survival of spores similar to that observed with RUV.

The lamp is equipped with a shutter to study the effects at relatively short exposure times. Figure 2d shows an inactivation curve with samples of strain WT 168 irradiated at 10 K and vacuum. Almost no cells can survive fluences $\geq 60 \text{ kJ m}^{-2}$ using the full spectrum (vacuum UV + far UV) of the H_2 lamp which corresponds in Fig. 2d to $\geq 10 \text{ kJ m}^{-2}$ with the vacuum UV.

To explain the inactivation profiles induced by vacuum UV irradiation alone, we checked for the biological effectiveness of this source at several wavelengths in the vacuum UV using interference filters of centre wavelengths 1,600, 1,400 and 1,200 Å (half-width of filters ~ 200 Å).

The radiation transmitted through these three filters was measured by CO_2 dosimetry. Note that the hydrogen lamp spectral output, particularly at the very short wavelengths, is unstable so that the relative fluxes at 1,200 and 1,600 Å may vary by as much as a factor of 2 from run to run, however, this does not affect our results significantly. We examined survival of our samples at doses between 1 and 120 kJ m^{-2} . There was no significant destruction of spores when irradiating with wavelengths $\lambda = 1,400$ and 1,215 Å, within our chosen dose rates, at both room temperature and 10 K. On the other hand, we observed significant inactivation of spores when irradiating with $\lambda = 1,600$ Å light in room-temperature conditions.

Irradiation with the total spectrum (vacuum UV + far UV) shows that spores are killed very effectively but significantly less at 10 K than at room temperature (Fig. 2e, f). Substitution of the MgF_2 window (which is transparent down to 1,100 Å) by a quartz window (which only transmits radiation longward of 2,000 Å) leads to the result that the remaining far UV spectrum of the lamp is as effective as the total spectrum at both temperature extremes (Fig. 3). We find that irradiation with wavelengths $\lambda > 2,000$ Å accounts for essentially all the inactivation when comparison is made with the inactivation curve of the H_2 lamp's full spectrum.

Thus, we conclude that: (1) wavelengths $\leq 1,400$ Å do not affect the spores within our chosen dose rates; (2) 1,600-Å light inactivates spores but it takes doses delivered at room temperature, which are significantly higher than applying the lamp full spectrum; (3) the total flux of the lamp inactivates at doses $\geq 60 \text{ kJ m}^{-2}$ (vacuum UV $> 10 \text{ kJ m}^{-2}$). It is evident that photons of energies (E) < 6 eV are the more effective ones in spore killing than those with $E > 6$ eV.

One factor in the interstellar medium that may provide a mechanism for the protection of spores is that of absorption of

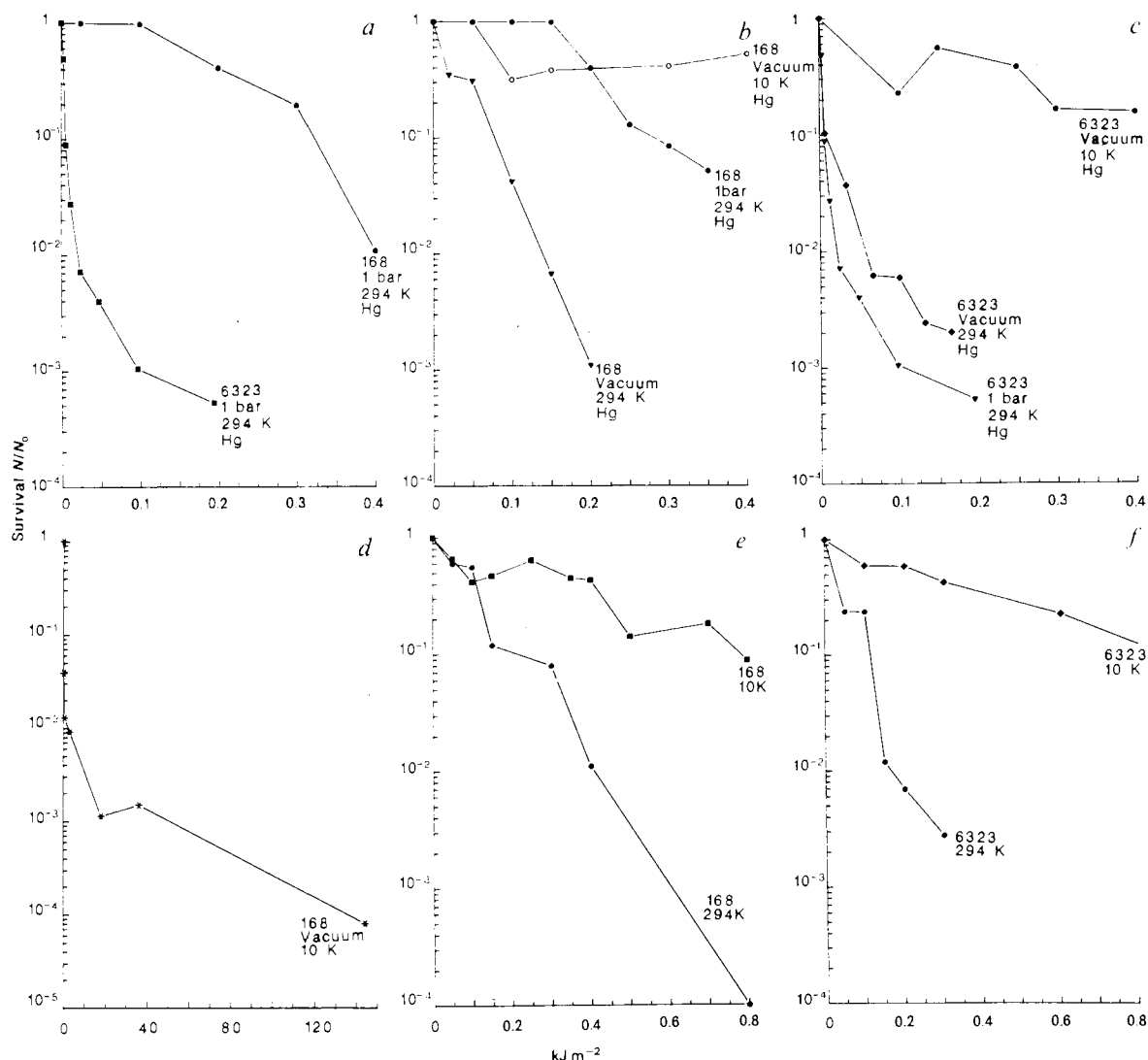


Fig. 2 Semilogarithmic plots of inactivation curves for *Bacillus subtilis* spores. *a*, WT 168 and TKJ 6323 spores irradiated at room temperature and 1 bar with a mercury low-pressure lamp. *b*, WT 168 spores irradiated at 1 bar and room temperature. Irradiations in vacuum are performed at room temperature and 10 K. Air-dried spores are irradiated at 1 bar and room temperature. *c*, TKJ 6323 spores irradiated in vacuum at room temperature and 10 K. For comparison air-dried spores are irradiated at 1 bar and room temperature. *d*, WT 168 spores irradiated with a H_2 lamp at 10 K in vacuum with high doses. Full spectrum vacuum UV+far UV. Abscissa scale is vacuum UV alone. *e*, WT 168 spores irradiated in vacuum at 10 K and room temperature with a H_2 lamp. Full spectrum vacuum UV+far UV. Abscissa scale is vacuum UV alone. *f*, TKJ 6323 spores, irradiations as in *e*.

UV light within mantles of molecules which may grow to a thickness of 1 μm within the mean lifetime of a molecular cloud.

We therefore created an artificial mantle on spores to check whether this influences the survival in simulated interstellar conditions. We deposited a mixture of $H_2O:CH_4:NH_3:CO$ (1:1:1:1) on top of the spores on a cold finger at 10 K and vacuum ($<1 \times 10^{-6}$ torr). The mantle thickness is measured by monitoring the interference pattern of light of a He/Ne-laser by means of a photomultiplier tube. Mantles of various thicknesses were deposited and the system irradiated with the H_2 lamp.

We find that such mantles do indeed protect spores from vacuum UV irradiation but not from far UV. The reason for this is that these ices consist of simple molecules which absorb strongly in the vacuum UV but are essentially transparent in the far UV. Thus using the full H_2 -lamp spectrum on $\geq 0.5\text{-}\mu\text{m}$ mantle-coated spores, at 10 K, we find, as expected, that the sensitivity of TKJ 6323 is not noticeably decreased whereas that of WT 168 is decreased by a factor of 10. This demonstrates again the fact that the vacuum UV damages are distinguishably different from those due to far UV.

Note that the mantles which accrete in interstellar space have been themselves irradiated by the UV and consequently consist

of complex rather than simple molecules. Such photo-processed mantle materials should absorb strongly at wavelengths $<3,000 \text{ \AA}$ and consequently provide effective shielding in the far as well as in the vacuum UV.

Discussion

It seems possible to distinguish between the types of damage occurring within and outside the core of a spore, as a function of wavelength. The outer layers of the spore protect the core by attenuating UV light (see ref. 19 for review). Within the core, damage may be divided into several types: (1) cyclobutane-type dimers; (2) 5-thymine-5,6-dihydrothymine spore photoproduct (TDHT); (3) DNA-protein crosslinks; (4) DNA strand breaks; (5) DNA adducts. Outside the core, damages are produced within proteins rather than in DNA.

In general, there is an energy discrimination between the various types of damage such that, for example, TDHT forms as a result of low-energy photons ($<6 \text{ eV}$) whereas the DNA-protein crosslink and strand break cross-sections are peaked at $\sim 10 \text{ eV}$ (ref. 20).

Because at extremely low temperatures and ultra-high vacuum we are dealing with very dry states of the system, irradiation mainly induces radicals leading to TDHT through a well-established

Table 2 Stain exposure conditions

		T		F_{10}^*	F_{10} (sample)	D_{37}	Absorption
	P	(K)	UV	(J m ⁻²)	F_{10} (reference)	(linear regression) (J m ⁻²)	cross-section σ (m ² photon ⁻¹)
168	1 bar	294	Hg lamp	325	1(ref)	170	4.58×10^{-21}
168	vacuum	294	Hg lamp	75	0.23	39	2.0×10^{-20}
168	vacuum	10	Hg lamp	—	$\gg 1$	—	$\leq 2 \times 10^{-23}$
168	vacuum	294	H ₂ lamp	200	0.615	105	1.18×10^{-20}
168	vacuum	10	H ₂ lamp	800	2.46	220	5.65×10^{-21}
6323	1 bar	294	Hg lamp	35	0.1	12.7	6.14×10^{-20}
6323	vacuum	294	Hg lamp	32	0.09	28.7	2.71×10^{-20}
6323	vacuum	10	Hg lamp	—	$\gg 1$	—	$\leq 1 \times 10^{-22}$
6323	vacuum	294	H ₂ lamp	200	0.285	100	1.24×10^{-20}
6323	vacuum	10	H ₂ lamp	950	2.71	180	6.9×10^{-21}

* These fluences for the H₂ lamp are calculated using only the range 1,200–2,000 Å. The fluences including the far UV are six times larger.

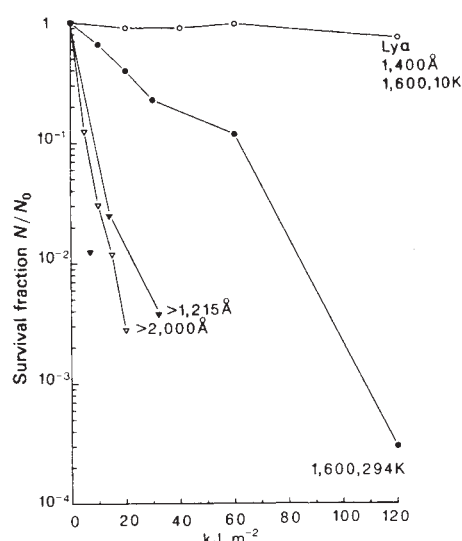


Fig. 3 WT 168 spores irradiated with a H₂ lamp in vacuum at various wavelengths. The fluences, in each case, are determined by the specified wavelength region of irradiation.

lished mechanism²¹. Note, therefore, that we are not inducing singlet or triplet structure. At 10 K we observed no discrimination between spores of strains WT 168 and TKJ 6323 irradiated with RUV. This immediately suggests that the UV photolysis induced at low temperatures is repair independent or, in other words, that other photoproducts which are not amenable to any known cellular repair mechanism are produced.

DNA-protein crosslinking is increased when spores are RUV-irradiated in vacuum²². The enhancement for the formation of this photoproduct in vacuum is about a factor of 12 greater than irradiations at 1 bar and correlates with an increase in inactivation²³.

The increase of non-DNA damage has been described by Smith and O'Leary²⁴. They point out that UV-induced photoproducts which are not amenable to photoreactivation and dark repair may be formed at considerably lower rates than other types of photochemical lesions but may, nevertheless, be more deleterious to cells²⁵.

The chemical analysis of DNA photoproducts induced in spores by vacuum irradiation of solar UV (to be published elsewhere) indicates that TDHT is the major photoproduct of DNA formed at both 10 K and room temperature. Thus, in accordance with our survival data, TDHT does not contribute very much to inactivation of spores when irradiated at 10 K.

This observation is consistent with a result of Ashwood-Smith *et al.*²⁶ who found that the lack of supersensitivity of purified DNA to RUV irradiation at moderately low temperatures (–79 °C) strongly suggests that the lesions responsible for the low-temperature supersensitivity are not primarily associated with changes in DNA. For phages, it has been shown with dose

rates comparable to those in our experiment, that crosslinks between DNA and proteins are induced through long-lived radicals. In the dry spore, this mechanism might also be plausible and is probably suppressed at extremely low temperatures because of inhibited diffusion^{27,28}.

When using the H₂ lamp we expect a shift in the target chromophore. This light source emits photons of much higher energies than the RUV source. Although these energies may split intramolecular bonds of DNA and peptides^{29,30}, and the Lyman α of the lamp coincides with the maximum efficiency for DNA strand breaking²⁰, nevertheless, we observe that most of the effective damage arises from the far UV rather than the vacuum UV.

Thus, at 10 K, we found, surprisingly, that the vacuum UV and RUV kill much fewer cells than the residual spectrum, which is that between 2,000 and 3,000 Å exclusive of the 2,537-Å Hg line. Further spectral studies are needed to determine the precise region of the susceptibilities at 10 K.

Astrophysical implications

According to our survival data, a dose of ~6 kJ m⁻² of vacuum and far UV inactivates spores to F_{10} at 10 K. Noting the fact that the H₂ lamp emits five times as much energy in the far UV relative to the vacuum UV as compared with the interstellar medium (ISM) implies that F_{10} for the diffuse interstellar medium occurs in ~150 yr.

To inactivate spores to $F_{0.1}$ (99.9% inactivated) requires a dose of 120 kJ m⁻² (20 kJ m⁻² of vacuum UV in Fig. 2d), which corresponds to 2,500 yr of irradiation in the diffuse ISM where the UV radiation is not attenuated. Spores exposed to such a dose can be regarded as dead if we accept the theory that the levelling off of the inactivation curve is an artefact of our experimental method³¹.

Two possibilities exist, which may lead to the prolongation of the mean lifetime of a bacillus in the ISM: (1) spores are deposited within dense clouds in which the UV radiation is attenuated by several orders of magnitude; (2) spores accrete mantles of condensable matter which reduces the penetration of UV to the outer layer of the spore by a factor as large as 10 or more.

Both the far and vacuum UV in the centre of a cloud of density $n_H = 10^4$ cm⁻³ and radius of 1 pc are attenuated by a factor of 10^8 – 10^9 (ref. 32). Because there generally exist internal sources of UV in addition to the attenuated radiation, which comes from distant stars, it is more conservative to consider the mean UV radiation within a molecular cloud region to be reduced to not less than ~ 10^{-4} times that of the diffuse ISM radiation. This minimum is internally produced by conversion of cosmic-ray energy and stellar-wind energy^{33,34}.

The growth of a mantle on a bacterium occurs within a cloud by accretion of molecules at a rate³⁵:

$$\frac{da}{dt} = 3.43 \times 10^{-22} n_H \text{ cm s}^{-1}$$

Thus, in 1.5×10^5 yr, one should accrete a mantle of $\sim 0.15 \mu\text{m}$. This material is irradiated by UV during condensation, and, because of the production of new molecules and radicals, absorbs in the UV more than the originally simple molecules. The attenuation of $3,000\text{-}\text{\AA}$ radiation (and shorter) through such a layer is greater than a factor of 30 as measured in our laboratory.

Combining the effects of accretion with cloud attenuation leads to a biological time (F_{10}) of 4.5–45 Myr, which seems to be long enough for spores to be transported from one solar system to another.

Conclusions

We have presented experimental evidence for the effects of very low temperature and UV radiation, characteristic of the ISM, on the survival of bacteria. In the most general environment in space, 10% survival times are only of the order of hundreds of years, too short for panspermia to work.

However, in a substantial fraction of space within dark clouds, we have shown that, even with conservative figures, survival times as long as millions to tens of millions of years are attainable. In such conditions, clouds could transport organisms from one solar system to another in times significantly shorter than the mean survival time. This occurs with significant probability. However, the problem of survival of bacteria during the ejection and the ultimate deposition on a non-hostile host planet still has not been addressed.

Received 21 December 1984; accepted 22 May 1985.

- Schidlowsky, U., Appel, P. W. V., Eichman, R. & Junge, C. E. *Geochim. cosmochim. Acta* **43**, 189–199 (1979).
- Brooks, J. & Shaw, G. *Origin and Development of Living Systems*, 354 (Academic, New York, 1973).
- Kven Volden, K. A., Lawless, J. G. & Ponnampuruma, C. *Proc. natn. Acad. Sci. U.S.A.* **68**, 486–490 (1971).
- Greenberg, J. M. *Origins of Life* **14**, 25–36 (1984).
- Hoyle, F. & Wickramasinghe, N. C. *New Scientist* 412–415 (13 August 1981).
- Oparin, A. I. *The Origins of Life* (Dover, New York, 1953).
- Blome, H. J. & Priester, W. *Naturwissenschaften* **71**, 456–467 (1984).
- Eigen, M. *Naturwissenschaften* **58**, 465–523 (1971).
- Eigen, M. & Schuster, F. *The Hypercycle—a Principle of Natural Self-Organisation* (Springer, New York, 1979).
- Arrhenius, S. in *The Quest for Extraterrestrial Life* (ed. Goldsmith, D.) 32–33 (University Science Books, California, 1980) [transl. by Goldsmith, D. *die Umschau* **7**, 481 (1903)].
- Spitzer, L. Jr *Physical Processes in the Interstellar Medium*, 227 (Wiley Interscience, New York, 1978).
- Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
- Greenberg, J. M. *Astr. Astrophys.* **12**, 240 (1971).
- Greenberg, J. M. & Shah, G. A. *Astr. Astrophys.* **12**, 250–257 (1971).
- Greenberg, J. M. *Scient. Am.* **250** (1984).

The only way a bacterium can survive above the Earth's atmosphere is if it is coated by a mantle of some material which attenuates the solar UV radiation by a factor of $\sim 10^9$. A thickness of $0.9 \mu\text{m}$ of material with an imaginary index of refraction of 0.5 would be sufficient, but how a spore can come to be so coated and blown out to space remains completely unknown. A possible suggestion would be the ejection to the upper atmosphere accompanying some explosive event—perhaps a comet or meteorite collision. The ejection process may be non-destructive if the impacting object clears the atmosphere and leaves a channel for the escape of high-speed ejecta³⁶. However, one still requires that coincident with the meteorite/comet collision there occurs the passage of the solar system through a dense molecular cloud which can capture the ejecta. The complementary question of deposition is answered affirmatively if, when the spore enters a new solar system, it already has a mantle of absorbing material of the required thickness.

Although we have not completely answered the question of panspermia, we have provided some experimental results which lead to well-defined restrictions on the required processes.

This work was supported in part by NASA grant NGR-33-018-148. P.W. thanks the European Space Agency for a fellowship during the tenure of this research. We thank G. Mohn of Leiden University for allowing us to use some of his laboratory facilities, G. Horneck of the Deutsche Forschungsgemeinschaft für Luft- und Raumfahrt for helpful advice, and F. Baas for his invaluable advice and assistance.

- Hagen, W., Allamandola, L. J. & Greenberg, J. M. *Astrophys. Space Sci.* **65**, 215–240 (1979).
- Horneck, G. *Adv. Space Res.* **1**, 39–48 (1981).
- Ashwood-Smith, M. J., Copeland, J. & Wilcockson, J. *Nature* **217**, 337 (1968).
- Ito, T., Ito, A., Hieda, K. & Kabayashi, K. *Radiat. Res.* **96**, 532–548 (1983).
- Wirths, A. & Jung, A. *Photochem. Photobiol.* **15**, 325–330 (1972).
- Giese, A. C. (ed.) *Photophysiology* Vol. 7, 207 (Academic, New York, 1972).
- Bücker, H. & Horneck, G. *Life Sci. Space Res.* **12**, 209–213 (1974).
- Horneck, G. et al. *Proc. 2nd European Symp. Life Sci. Res. Space* (ESA Sp-212, 1984).
- Smith, K. C. & O'Leary, M. E. *Science* **155**, 1024 (1967).
- Simukova, N. A. & Budowsky, E. I. *FEBS Lett.* **38**, 299–303 (1974).
- Ashwood-Smith, M. H. & Bridges, B. A. *Proc. R. Soc.* **168B**, 194–202 (1967).
- Fekete, A. & Rontó, G. Y. *Stud. biophys.* **80**, 165–171 (1980).
- McIntosh, D., Moskovito, M. & Ozin, G. A. *Inorg. Chem.* **15**, 1669 (1976).
- Dodonova, N. Y., Kiseleva, U. N., Tsyganenko, N. M. & Remisova, L. A. *Photochem. Photobiol.* **35**, 129–132 (1982).
- Iwanami, S. & Oda, N. *Radiat. Res.* **90**, 466–478 (1982).
- Casolari, A. J. *theor. Biol.* **88**, 1–34 (1981).
- Chliewicki, G. & Greenberg, J. M. *Mon. Not. R. astr. Soc.* **211**, 719–730 (1981).
- Prasad, S. S. & Tarafdar, S. P. *Astrophys. J.* **267**, 603 (1983).
- Silk, J. & Norman, C. *IAU Symp.* **87**, 165–172 (1980).
- Greenberg, J. M. in *Cosmic Dust* (ed. McDonnell, J. A. M.) 187–294 (Wiley Interscience, New York, 1978).
- Melosh, H. J. *Geology* (in the press).

A peculiar supernova in the spiral galaxy NGC4618

Alexei V. Filippenko*† & Wallace L. W. Sargent†

* Department of Astronomy, University of California, Berkeley, California 94720, USA
† Palomar Observatory, California Institute of Technology, Pasadena, California 91125, USA

Optical spectra of a bright stellar object near the nucleus of the spiral galaxy NGC4618 reveal strong, very broad emission lines similar to those in quasars but having the wrong relative wavelengths. Although lines of hydrogen and helium are absent, the most prominent features can be attributed to neutral atoms of oxygen, sodium, and magnesium at the redshift of NGC4618. The object is almost certainly a supernova whose highly unusual spectrum may be indicative of a fundamentally new subclass.

DURING an extensive spectroscopic survey of nearby galactic nuclei¹, we have discovered² a bright starlike object in the peculiar SBbc(rs)II.2 spiral NGC4618 (brightness $B_T = 11.20$ mag, velocity $v = 532 \text{ km s}^{-1}$, $b^\text{II} = 75^\circ 8'$)³, also known as Arp 23 and VV73. It appeared to be close to one end of a prominent bar (position angle, $PA \approx 64^\circ$) in the central region

of the galaxy, and was assumed to be the nucleus. Its spectrum is unlike that of any astronomical object yet published.

Observations

Spectra of NGC4618 were obtained on 28.52 February 1985 UT with the double spectrograph⁴ at the Cassegrain focus of the

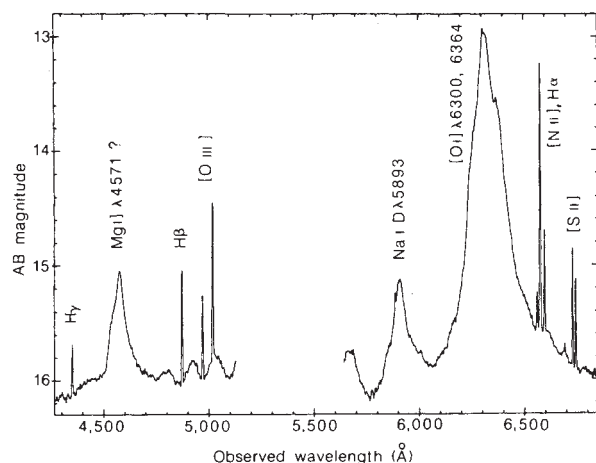


Fig. 1 CCD spectra of the starlike object and its surrounding H II region near the nucleus of NGC4618. Ordinate units are given by $AB = -2.5 \log(f_\nu) - 48.6$, where f_ν is in $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$, but the blue spectrum is probably ~ 0.1 mag fainter than displayed here. All narrow lines are associated with the H II region, while the broad features, which strongly resemble those in quasars, are almost certainly produced by a supernova in NGC4618. The data have been smoothed with a gaussian ($3 \text{ \AA}$ FWHM), making the peak intensities of narrow lines appear smaller than in the original spectra.

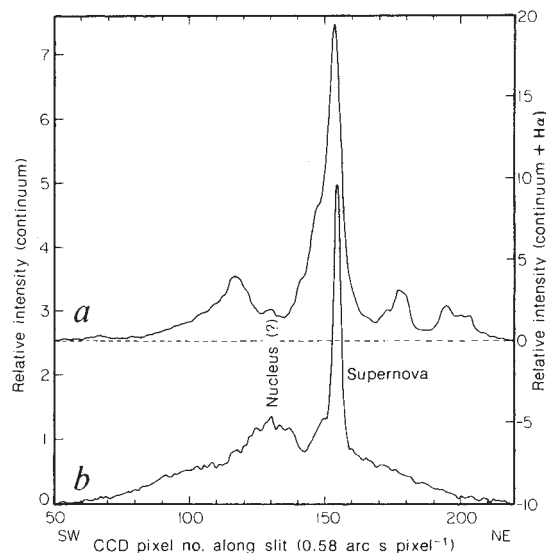


Fig. 2 Plots of the average continuum ($6,608\text{--}6,632 \text{ \AA}$) emission (b) and average continuum plus $H\alpha$ ($6,570\text{--}6,578 \text{ \AA}$) emission (a) along the central bar in NGC4618. The apparent nucleus, as defined by the continuum, is ~ 14 arc s from the supernova, although the actual nucleus may be hidden by a thick dust lane near pixel 142. Many H II regions are visible, in (a), the most intense of which surrounds the supernova and is brightest on the south-west side.

Hale 5.08-m telescope at Palomar Observatory. The slit (~ 2 arc min in length) was placed along the conspicuous central bar and included the stellar object. A dichroic filter directly behind the slit reflected blue light ($\lambda \leq 5,500 \text{ \AA}$) to one camera and transmitted red light ($\lambda \geq 5,500 \text{ \AA}$) to the other. Detectors consisted of two charge-coupled devices (CCDs) (Texas Instruments), each with 800×800 square pixels. Gratings having 1,200 and 600 grooves mm^{-1} were used in the first order to achieve resolutions (full width at half-maximum, FWHM) of $\sim 2.6 \text{ \AA}$ and $\sim 4.3 \text{ \AA}$ in the red and blue cameras (respectively) with a slit width of 2 arc s. The first integration (20 min) covered the intervals $\lambda\lambda 4,250\text{--}5,125$ and $\lambda\lambda 6,200\text{--}6,850$, whereas the second (15 min) provided useful data over $\lambda\lambda 3,680\text{--}4,120$ and $\lambda\lambda 5,620\text{--}6,280$. The seeing was 1.5–2 arc s at all times.

Normal procedures were used to calibrate the spectra. A bias level was subtracted from each data frame, and division by the featureless spectrum of a hot tungsten lamp removed local variations in the sensitivity of the CCDs. Pixels affected by cosmic rays were eliminated by inspection. Wavelengths were determined by fitting cubic polynomials to unblended emission lines of Fe, Ar and Ne in spectra of comparison arcs. Small distortions of the spectra both parallel and perpendicular to the dispersion were similarly corrected. Contamination from the background sky was subtracted by using regions of the slit far from the galactic nucleus. Bright standard stars⁵ served to calibrate the broadband instrumental response as a function of wavelength. Although NGC4618 was observed in nearly photometric conditions, there is a relative uncertainty of ≤ 0.1 mag between the red and blue halves because of an instability in the dichroic. Differential losses from atmospheric dispersion⁶ were negligible in all but the bluest spectrum ($\lambda\lambda 3,680\text{--}4,120$).

Results

Figure 1 illustrates the spectrum of the stellar object in NGC4618. It consists of intense, broad emission lines probably superposed on a smooth continuum. Several weaker, but nevertheless broad, features are also likely to be emission lines. In addition, there are many narrow lines resembling those in H II regions ionized by hot stars. The spectral interval $\lambda\lambda 3,680\text{--}4,120$ is not plotted because it is of relatively poor quality and exhibits only a single narrow emission feature at $\sim 3,734 \text{ \AA}$; broad lines, if present, must have very low contrast above the continuum.

Inspection of the long-slit CCD spectra reveals a series of H II regions which are presumably responsible for the mottled appearance of the central bar in NGC4618 (Fig. 2). The brightest of these is almost coincident with the starlike object; its peak is displaced by at most ~ 1 arc s, while its centroid is a few arc seconds away. In the top portion of Fig. 2 it appears as a 'shoulder' predominantly on one side (south-west) of the unresolved spike. Because the effective entrance aperture for the spectrum in Fig. 1 is 2×8.7 arc s, much of the light from this H II region is included and it obviously accounts for all the narrow emission lines. Their relative intensities do not differ substantially from those in other ionized clouds along the slit. The ratio $I([\text{S II}] \lambda 6,716)/I([\text{S II}] \lambda 6,731) \approx 1.40$, indicating⁷ that the electron density (n_e) is $\leq 60 \text{ cm}^{-3}$. Assuming normal Case B recombination⁸, a visual extinction of ~ 1.0 mag is implied by the observed strengths of $H\beta$ and $H\gamma$. This may or may not apply to the stellar object, depending on its actual spatial location with respect to the H II region. The narrow line at $\sim 3,734 \text{ \AA}$ mentioned earlier is clearly $[\text{O II}] \lambda 3,727$.

Intensities, equivalent widths, wavelengths, and profile widths (FWHM; full width at zero-intensity, FWZI) of the emission lines are given in Table 1. The narrow features from the H II region are unresolved in all of the spectra, so their widths are not tabulated. When the narrow lines appear on top of broad emission features, their equivalent widths are calculated with respect to the estimated continuum. Most measurements of broad lines (other than wavelengths) are seriously affected by the choice of continuum level, and, therefore, have considerable uncertainties. It is possible, for example, that the intrinsic continuum in the peculiar object is extremely weak if the wings of the emission lines are very extended ($\text{FWZI} \geq 700\text{--}800 \text{ \AA}$), as may be the case for the feature at $\sim 6,320 \text{ \AA}$ (Fig. 1). If so, the strengths and widths (particularly the FWZI) of broad lines in Table 1 are gross underestimates. Note that the FWZI, rather than a more well-defined quantity such as the full width at 10% intensity, is listed to emphasize the enormous maximum gas velocities which are derived under the assumption of Doppler broadening.

Although the widths and profiles of the broad emission lines resemble those commonly observed in the spectra of quasars (QSOs), their relative wavelengths do not, even if large blueshifts are allowed. Moreover, the object cannot be a high-redshift QSO gravitationally lensed by the nucleus of NGC4618, unlike the

Table 1 Measurements of emission lines

Line	Observed λ of peak (Å)	Flux (10^{-14} erg s $^{-1}$ cm $^{-2}$)*	EW (Å)†	FWHM (Å)‡	FWZI (Å)§
H γ λ 4,340	4,349.0	0.64	3.1	—	—
Mg I] λ 4,571?	4,575	21	100	55	270
λ 4,786	4,795	1.2	6.2	62	110
H β λ 4,861	4,870.3	1.59	8.89	—	—
[Ca I]? λ 4,912	4,921	1.6	9.1	53	86
[O III] λ 4,959	4,968.3	0.976	5.53	—	—
[O III] λ 5,007	5,016.3	2.91	16.6	—	—
Fe II? λ 5,022	5,031	5.0	29	64	120
λ 5,659	5,669	>4.6¶	>40¶	>70¶	>140¶
He I λ 5,876	5,886.2	0.16	1.5	—	—
Na I λ 5,893	5,905	17	160	69	290
[O I] λ 6,300	6,312#	220	2100	110	700
[O I] λ 6,364	6,375#				
[N II] λ 6,548	6,560.0	0.25	2.6	—	—
H α λ 6,563	6,574.8	5.26	53.3	—	—
[N II] λ 6,583	6,595.5	0.82	8.4	—	—
λ 6,600	6,612:	1:	10:	40:	80:
λ 6,677	6,689	0.3:	3:	20:	30:
He I λ 6,678	6,690.5	0.05:	0.5:	—	—
[S II] λ 6,716	6,728.7	0.71	7.3	—	—
[S II] λ 6,731	6,743.1	0.51	5.2	—	—

Note that some values differ from preliminary ones (ref. 2). Colon means very uncertain.

* Not corrected for extinction in NGC4618. Galactic $A_v \approx 0$.

† Equivalent width. Placement of continuum highly uncertain.

‡ Full width at half-maximum. Unresolved lines have no entry (H II region).

§ Full width at zero-intensity. Unresolved lines have no entry (H II region).

|| Derived from $z = 0.0018$ and the observed λ ; exact identity unknown.

¶ Lower limit; end of spectrum truncates emission line (see Fig. 1).

Approximate centroid (rather than peak). Peak slightly blueshifted.

recently discovered quasar⁹ within 0.3 arc s of the nucleus of the foreground spiral galaxy 2237+0305. Figure 2 shows that the apparent nucleus of NGC4618 is displaced from the peculiar object by ~ 14 arc s, and the galaxy is neither distant enough nor sufficiently centrally concentrated to produce such a large separation. Lensing by the fortuitous alignment of a massive black hole outside the nucleus is also highly unlikely.

The great widths of the most intense features (FWHM $\approx 3,000$ – $5,000$ km s $^{-1}$; FWZI $\approx 15,000$ – $30,000$ km s $^{-1}$) suggest that the object is some kind of explosive event. As will soon be shown, physical association with NGC4618 ($z = 0.0018$) is likely; the object's absolute magnitude is, therefore, $M_v \approx -14$ (or even -15 if the visual extinction $A_v \approx 1.0$ mag), much too bright for a nova. The remaining possibility among objects of known type is a supernova (SN)—although its spectrum differs markedly from that of any supernova yet published. The transient nature of the object is evident because nothing peculiar has previously been noted in spectra of NGC4618 obtained during galaxy redshift surveys^{10,11}. Furthermore, the central 'bar' was mentioned in each case, but there was no indication of the presence of an extremely bright, starlike 'nucleus' (which would be unusual in a galaxy whose overall appearance is fairly irregular). Solid proof, however, is that the object appears on photographic plates taken at Lick Observatory on 16 March 1985 UT, whereas a similar Lick plate exposed on 26 May 1914 UT shows only a bright H II region at the same location¹². The position (equinox 1950.0) of the 'star' was accurately measured from a plate taken with the 1.2-m Palomar Schmidt telescope¹³: $\alpha = 12$ h 39 min 09.56 s, $\delta = +41^\circ 25' 32''.7$.

A partial spectrum of the true nucleus of NGC4618 (or, more precisely, of the position along the central bar which has the highest continuum luminosity) is illustrated in Fig. 3. Prominent, relatively broad Balmer absorption lines indicate the presence of hot, massive young stars, showing that active star formation has recently occurred. Emission lines of H α , H β , [O II] λ 3,727, [O III] λ 4,959, 5,007, [N II] λ 6,548, 6,583, and [S II] λ 6,716, 6,731 are also visible. The H II regions are even more pronounced elsewhere along the slit (see Fig. 2), especially very near the stellar object. Thus, the environment in the central

regions of NGC4618 is exactly that in which supernovae are expected to be found.

Figures 4 and 5 show detailed portions of the red and blue regions of the supernova's spectrum. The profile of the strongest broad emission line (Fig. 4), whose equivalent width is at least $\sim 2,100$ Å, is asymmetrical and has considerable structure. Three relatively narrow peaks are evident near $\sim 6,312$ Å, but exactly the same pattern also appears at a weaker level near $\sim 6,375$ Å. Analysis of the profile demonstrates that the features are undoubtedly the [O I] λ 6,300, 6,364 lines at the redshift ($z = 0.0018$) of NGC4618. The reasonably certain identification of the strong line at $\sim 5,905$ Å (Fig. 1) with Na I D λ 5,890, 5,896 supports this conclusion, as does the less firm association of the feature at $\sim 4,575$ Å (Fig. 5) with Mg I] λ 4,571. Weaker broad lines are visible at wavelengths of $\sim 4,795$, 4,921, 5,031 and 5,669 Å, and possibly also at $\sim 6,612$ and 6,689 Å. These remaining lines are unidentified, but the species responsible for them may include Ca⁰, Fe⁰, and Fe⁺ (for example, Fe II λ 5,018, [Ca I] λ 4,913, 4,916). Thus, the object probably has a low-excitation spectrum. In particular, the three most prominent features represent the least energetic resonance transitions of neutral O, Na, and Mg; the corresponding upper levels are all easily populated by collisions, even at low temperatures.

Broad lines of H and He are undoubtedly absent, except perhaps for the very weak feature at $\sim 6,689$ Å which might be He I λ 6,678. This is improbable, however, as other lines of He (such as λ 3,889, λ 4,471, λ 5,876) should also be visible. The strong line at $\sim 5,905$ Å is definitely not He I λ 5,876 unless it is redshifted by 0.0049 rather than 0.0018. Emission lines of Na I D, on the other hand, have previously been seen in the spectra of novae, T Tauri stars, and various early-type stars¹⁴, so the identification is almost certain. Note that the lack of H and He lines does not necessarily imply a severe deficiency of these elements in the emitting gas, because there is no evidence for an ionizing continuum, and the electron temperature (T_e) may be too low for effective collisional excitation of even the first upper states.

A measured flux of $\sim 2.2 \times 10^{-12}$ erg s $^{-1}$ cm $^{-2}$ for the [O I] λ 6,300, 6,364 lines implies that the supernova contains at least

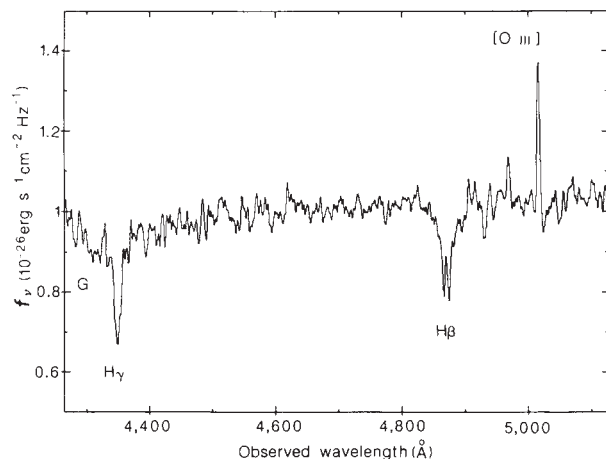


Fig. 3 Spectrum of the apparent nucleus of NGC4618, as marked in Fig. 2. Note the strong, relatively broad Balmer absorption lines indicative of hot young stars. $H\beta$ emission is visible within the corresponding absorption. Other emission lines from photoionized gas are also present, and dominate the spectrum in the red region ($\lambda\lambda 6,200$ – $6,850$).

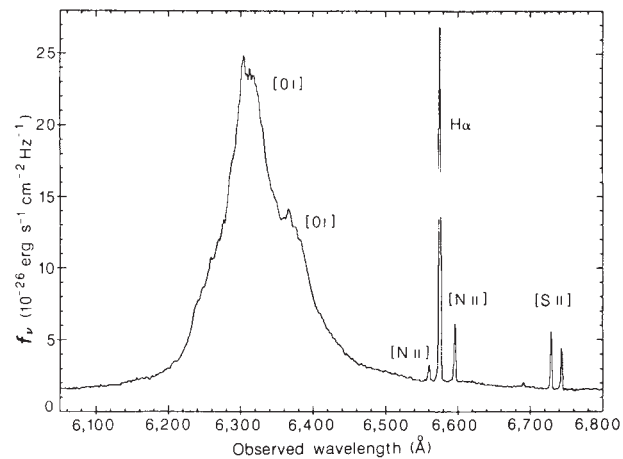


Fig. 4 Small, unsmoothed section of the spectrum in Fig. 1, showing the broad [O I] doublet in detail. [O I] $\lambda 6,364$ is roughly one-third the intensity of [O I] $\lambda 6,300$, and both lines have very similar complex structure. Note the absence of broad $H\alpha$ emission. Narrow lines are produced by low-density gas ($n_e \leq 60 \text{ cm}^{-3}$) in an H II region.

$0.01 M_{\odot}$ of oxygen, and probably much more ($\sim 1 M_{\odot}$). The remarkable strength of the doublet, together with the absence of $H\alpha$, suggests that $n_e \leq 10^6 \text{ cm}^{-3}$ unless the gas is almost entirely devoid of hydrogen or T_e is very low. In any case, n_e cannot be much higher than $\sim 10^7 \text{ cm}^{-3}$, since [O I] $\lambda 5,577$ would then be sufficiently intense to produce a strong red wing near $\sim 5,625 \text{ \AA}$, the blue limit of the red CCD spectrum (Fig. 1). It is highly unlikely that the feature observed near this region corresponds to [O I] $\lambda 5,577$, as it appears to peak at $\sim 5,669 \text{ \AA}$. In fact, n_e is likely to be quite small; there is no ionized hydrogen, and most of the other elements, with the exception of calcium³³, seem to be primarily neutral.

The strong line at $\sim 4,575 \text{ \AA}$, tentatively attributed to Mg I] $\lambda 4,571$, merits special mention. Figure 5 reveals an asymmetrical profile with a shoulder on the blue side and an extended red wing, suggesting that the line may actually be a composite of several different features. If so, an alternative identification is with multiplet 2 of Si III, which has transitions at $\lambda\lambda 4,552.7$, $4,567.9$ and $4,574.8$. Although the average wavelength is $\sim 4,565 \text{ \AA}$, very close to the rest wavelength ($\sim 4,567 \text{ \AA}$) of the observed feature, Si III is unlikely to be the correct identification, as high-excitation lines of other elements should also be visible but are not. Similarly, if the emission resulted from multiplets 37 and 38 of Fe II, other Fe II transitions should also be more prominent. The intersystem Mg I] $\lambda 4,571$ line, on the other hand, need not be accompanied by additional Mg I features, and it is quite consistent with the presence of collisionally-excited [O I] $\lambda\lambda 6,300$, $6,364$ and Na I $\lambda 5,893$. In fact, Mg I] $\lambda 4,571$ is known to appear in the spectra of certain gaseous nebulae and Bep stars; it sometimes even dominates the blue-violet spectra of long-period variable stars¹⁴. In the NGC4618 object, a discrepancy of $\sim 4 \text{ \AA}$ between the measured wavelength of the Mg I] peak (redshift 0.0018 removed) and the laboratory value mimics that seen in the peaks of the [O I] doublet, although the centroids of the [O I] lines exhibit no such discrepancy (unlike Mg I]).

It is not clear whether the broad, wavy features in the blue half of the spectrum (Fig. 1) are absorption or emission lines, but we suspect they are the latter. Analysis of the Balmer decrement in the superposed H II region suggests that the blue spectrum in Fig. 1 is $\sim 0.1 \text{ mag}$ too bright relative to the red spectrum. If this calibration error exists then, after its removal, the valleys between the broad peaks in both halves can be joined together in a reasonably smooth manner to represent the continuum, and the peaks are probably emission lines. Data having greater spectral coverage are obviously needed to resolve this question. Moreover, these data will allow a search for other emission lines, some of which (such as [N I] $\lambda 5,199$, [C I]

$\lambda\lambda 9,823$, $9,849$) are expected from considerations of the present spectrum. The exact shape of the continuum will also be important to quantify.

Discussion

The long-term evolution of the spectra of Type I supernovae has been established mainly through studies^{15–17} of SN1937c in IC4182 and SN1972e in NGC5253. Before ~ 180 days after maximum light neither of these supernovae exhibited spectra anything like that observed in the NGC4618 object; they were dominated by blends of permitted Fe II absorption lines¹⁸. Beyond this epoch, however, the spectrum of SN1937c showed a pair of lines which Minkowski¹⁵ identified with the [O I] $\lambda\lambda 6,300$, $6,364$ doublet and which gradually came to dominate the red region. These were not observed in the spectrum of SN1972e (ref. 17). Furthermore, even at late stages the red emission lines in the spectrum of SN1937c were always weaker than features in the blue region, particularly a broad emission band at $\sim 4,650 \text{ \AA}$ whose wavelength differs greatly from the $\lambda 4,567$ (rest wavelength) feature in Fig. 5. Few Type I supernovae have been observed spectroscopically before maximum light¹⁹, and none resembled the object in NGC4618.

The spectra of Type II supernovae are more readily identifiable than those of Type I; a distinguishing feature is that Type II spectra always contain conspicuous Balmer emission lines. No Type II supernova has been as thoroughly studied as SN1937c and SN1972e. The object observed best was SN1970g in M101, which was followed intermittently for 443 days past maximum light¹⁶. Its spectrum showed emission lines of hydrogen over the entire period. This supernova, like two other Type II objects (SN1969l in NGC1058 and SN1973r in NGC3627), exhibited [O I] $\lambda\lambda 6,300$, $6,364$ emission ≥ 4 months past maximum, although these were never the strongest spectral features. In sharp contrast to the spectrum of the NGC4618 object, $H\alpha$ was by far the most prominent emission line in SN1979c (M100) after several months²⁰. Spectra obtained before and during the maximum of the recent Type II supernova in NGC4699 (whose progenitor was probably a Wolf-Rayet star)²¹ also do not resemble that in Fig. 1.

Zwicky²² classified several peculiar varieties of supernovae as Types III, IV, and V. It is not clear whether these are fundamentally different from Types I and II¹⁹, but, in any case, descriptions and microphotometer tracings of the spectra^{19,22} show no similarity to the supernova discussed here. Excellent atlases, such as that compiled by Greenstein and Minkowski²³, confirm these conclusions.

The estimated distance modulus of NGC4618 is $m - M = 30.1$ (assuming $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$). Thus, a Type I supernova

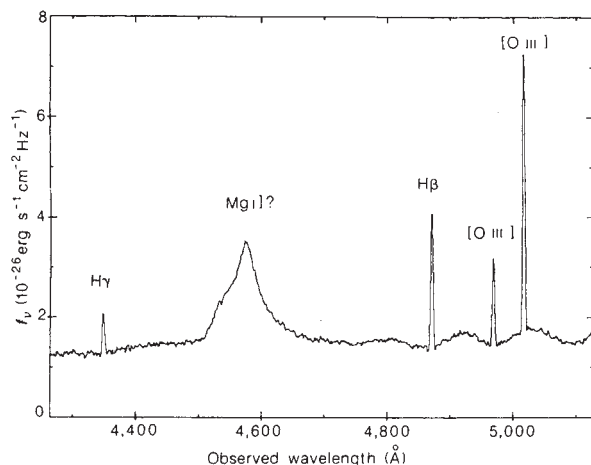


Fig. 5 Small, unsmoothed section of the spectrum in Fig. 1, showing the strong, broad, asymmetrical line attributed to Mg I λ 4,571. The profile is not identical to that of [O I] λ 6,300, although the peak is blueshifted relative to its expected position by about the same amount as the peaks of the [O I] lines (Fig. 4). Wavy features to the red of $\sim 4,700$ Å are probably emission lines, but their exact identity is unknown; possibilities include [Ca I] and Fe II. Narrow emission lines belong to the surrounding H II region.

with $M_v = -19.7$ at maximum light would have been an object of magnitude 10.4 in NGC4618; a Type II with $M_{\text{max}} = -19.0$ (ref. 24) would have been half as bright. The apparent visual magnitude ($m_v \approx 16$) at the end of February 1985 implies²⁵ that maximum light occurred ~ 160 days before discovery if the object was a typical Type I supernova or a moderately luminous Type II. A smaller time interval is obtained if the object was a fainter Type II or if the visual extinction (~ 1.0 mag) derived for the surrounding H II region also affects the supernova. The star could, therefore, have been at maximum in mid-September to late October 1984, when it was nearly at the same right ascension as the Sun and only $\sim 40^\circ$ north of it. If so, it may easily have been missed by optical searches, since a few months later it would have faded by several magnitudes and become more difficult to detect in the central region of NGC4618. The nucleus of NGC4618 is heavily overexposed, for example, in the National Geographic Society-Palomar Sky Survey plates. We emphasize, however, that based on the spectroscopic evidence there is no reason to associate the peculiar object in NGC4618 with any known kind of supernova at any phase.

Note that spectra of certain supernova remnants (SNRs) such as Cas A^{26,27}, N132D in the Large Magellanic Cloud²⁸, and the extraordinary object in NGC4449 (refs 29, 30) exhibit almost nothing but emission lines of oxygen. In particular, Balmer lines are completely absent. An intriguing possibility, but one which requires much more supporting evidence, is that the supernova in NGC4618 is a precursor to such SNRs. This might resolve the current uncertainty over whether Cas A resulted from a normal Type II supernova or from one in which the progenitor star lost its envelope and had a radius smaller than normal at the time of the explosion³¹. If the latter is true, then supernovae such as Cas A may be much fainter than $M_v \approx -19$ at maximum and hence are harder to detect. Indeed, the peculiar star in NGC4618 could have exploded only a short time (≤ 1 month) before its spectrum was obtained, so $M_v \approx -15$ to -16 (rather than -19) might be a reasonable estimate of its maximum luminosity.

Photographic plates obtained at the Sonneberg Observatory³² reveal that no conspicuous object brighter than $m_v \approx 12.5$ was present at the correct position between 28 February and 1 May 1984, on 30 November 1984, on 25 December 1984, and during the first two months of 1985. Although these data do not contradict our previous estimate of the supernova's date of birth (September or October 1984), they are also quite compatible

with the idea that this supernova, now designated SN1985f (ref. 32), exploded recently but was never very bright. Moreover, they eliminate the possibility that the object was a normal Type I or Type II supernova that exploded later than the beginning of November 1984 (or during March/April 1984). On the other hand, one cannot unambiguously exclude the possibility that this is a relatively old supernova which is fading very slowly and whose origins date back to 1983 or even earlier.

If the NGC4618 supernova really is similar to that which produced the Cas A remnant, then its progenitor must have evolved at least to the point where it created a substantial mass of oxygen. The circumstellar shell and ejecta from the explosion itself, however, must not yet have swept up an appreciable amount of interstellar gas. This is consistent with the probable youth of the supernova, and support is provided by the absence of radiation at radio wavelengths (6 cm, 20 cm) (R. A. Sramek, personal communication). Also, the lack of many strong lines from ionized species (or even high-excitation features of neutral atoms) means that the temperature of the emitting gas is much lower than in the SNRs mentioned above, which contain intense broad lines of [O II] and [O III] in addition to [O I]. Note that the SNR in NGC4449 is close to the centre of the galaxy (~ 60 arc s) and is embedded in an H II region, just like SN1985f near the nucleus of NGC4618.

Conclusions

We have discovered a probable supernova (SN1985f) ~ 14 arc s from the ill-defined nucleus of the spiral galaxy NGC4618. Its spectrum differs dramatically from those previously published for all other supernovae. In particular, a very strong (equivalent width $\geq 2,100$ Å), broad (FWHM $\geq 5,000$ km s⁻¹) emission line identified with the [O I] λ 6,300, 6,364 doublet dominates the spectrum, while lines of H and He appear to be completely absent. Positive and negative velocities of $\geq 15,000$ km s⁻¹ are indicated by the very extended wings of the profile. Other prominent emission lines include Na I D λ 5,893 and possibly Mg I λ 4,571. There are no obvious absorption lines, and the continuum is fairly weak.

If the maximum brightness and light curve of this supernova are similar to those of normal Type I and Type II objects, then we estimate that it achieved greatest brilliance ($m_v \approx 10.4$ mag) in mid-September or October 1984. Unfortunately, at that time NGC4618 was close to the Sun and could not be observed. A few months later the supernova may have been indistinguishable from the nucleus of the galaxy in deep photographs used for supernova searches (if such photographs of NGC4618 were even obtained), so its late detection is not surprising.

The peculiar spectral properties are reminiscent of certain oxygen-rich supernova remnants such as Cas A, N132D in the LMC, and the remnant in NGC4449. We speculate that the supernova in NGC4618 may represent a fundamentally new type, and that similar stars were progenitors of the anomalous remnants mentioned above. In this case, the star could have exploded only a short time before our observations were made, and its maximum luminosity was far below that of normal supernovae. Moreover, the spectral properties are expected to undergo drastic changes when a sufficiently large interaction with the interstellar medium or circumstellar shell finally occurs. It will be of interest to study the subsequent evolution of this bizarre object.

After this article was first submitted, many new spectra covering the entire optical window out to $1 \mu\text{m}$ were obtained at Lick and Palomar Observatories. Although a detailed analysis will be presented elsewhere³³, our major conclusions are confirmed. Strong absorption lines in the smooth continuum are definitely absent, but several broad emission lines appear in wavelength intervals outside those illustrated in Fig. 1. Some of these have already been reported¹²; note, however, that a very strong line initially identified as [O II] λ 7,325 is probably [Ca II] λ 7,308, while the faint feature between $\sim 4,000$ and $4,100$ Å is unlikely to be K I λ 4,046. The spectroscopic characteristics of the SN have shown few changes through mid-May 1985, and the overall

brightness seems to be diminishing slowly. Still unknown are the approximate birthday of this strange supernova, its maximum luminosity, and its relationship (if any) to remnants such as Cas A.

We thank J. Carrasco, J. Henning, and D. Tennant for assistance at Palomar Observatory. A photograph confirming the presence of the supernova was obtained by E. A. Harlan with

the 0.9-m Crossley telescope at Lick Observatory. R. A. Sramek provided radio data from the Very Large Array in Socorro, New Mexico. Informative discussions with K. Ebner, J. L. Greenstein, C. F. McKee, and J. B. Oke are appreciated. This work was supported by the Miller Institute for Basic Research in Science (University of California at Berkeley) and by NSF grant AST 82-16544.

Received 19 March; accepted 10 June 1985.

1. Filippenko, A. V. & Sargent, W. L. W. *Astrophys. J. Suppl.* **57**, 503-522 (1985).
2. Filippenko, A. V. & Sargent, W. L. W. *IAU Circ. No.* 4042 (1985).
3. Sandage, A. & Tammann, G. A. *A Revised Shapley-Ames Catalog of Bright Galaxies* (Carnegie Institution of Washington, 1981).
4. Oke, J. B. & Gunn, J. E. *Publ. astr. Soc. Pacif.* **94**, 586-594 (1982).
5. Oke, J. B. & Gunn, J. E. *Astrophys. J.* **266**, 713-717 (1983).
6. Filippenko, A. V. *Publ. astr. Soc. Pacif.* **94**, 715-721 (1982).
7. Cantó, J., Elliott, K. H., Meaburn, J. & Theokas, A. C. *Mon. Not. R. astr. Soc.* **193**, 911-919 (1980).
8. Brocklehurst, M. *Mon. Not. R. astr. Soc.* **153**, 471-490 (1971).
9. Huchra, J. et al. *Astr. J.* **90**, 691-696 (1985).
10. Humason, M. L., Mayall, N. U. & Sandage, A. R. *Astr. J.* **61**, 97-162 (1956).
11. Vaucouleurs, A. de & Vaucouleurs, G. de *Astr. J.* **72**, 730-737 (1967).
12. Filippenko, A. V. & Sargent, W. L. W. *IAU Circ. No.* 4048 (1985).
13. Gibson, J. & Schombert, J. *IAU Circ. No.* 4048 (1985).
14. Merrill, P. W. *Lines of the Chemical Elements in Astronomical Spectra* (Carnegie Institution of Washington, 1958).
15. Minkowski, R. *Astrophys. J.* **89**, 156-217 (1939).

16. Kirshner, R. P., Oke, J. B., Penston, M. V. & Searle, L. *Astrophys. J.* **185**, 303-322 (1973).
17. Kirshner, R. P. & Oke, J. B. *Astrophys. J.* **200**, 574-581 (1975).
18. Branch, D. et al. *Astrophys. J.* **270**, 123-139 (1983).
19. Oke, J. B. & Searle, L. A. *Rev. Astr. Astrophys.* **12**, 315-329 (1974).
20. Branch, D. et al. *Astrophys. J.* **244**, 780-804 (1981).
21. Niemela, V. S., Ruiz, M. T. & Phillips, M. M. *Astrophys. J.* **289**, 52-57 (1985).
22. Zwicky, F. in *Stars and Stellar Systems*, Vol. 8 (eds Aller, L. H. & McLaughlin, D. B.) 367-423 (University of Chicago Press, 1965).
23. Greenstein, J. L. & Minkowski, R. *Astrophys. J.* **182**, 225-243 (1973).
24. Tammann, G. A. in *Proc. NATO Advanced Study Institute on Supernovae* (eds Rees, M. J. & Stoneham, R. J.) 371-403 (Reidel, Dordrecht, 1981).
25. Shklovsky, I. S. *Supernovae* (Wiley, London, 1968).
26. Kirshner, R. P. & Chevalier, R. A. *Astrophys. J.* **218**, 142-147 (1977).
27. Chevalier, R. A. & Kirshner, R. P. *Astrophys. J.* **233**, 154-162 (1979).
28. Lasker, B. M. *Astrophys. J.* **223**, 109-121 (1978).
29. Balick, B. & Heckman, T. *Astrophys. J. Lett.* **226**, L7-L10 (1978).
30. Kirshner, R. P. & Blair, W. P. *Astrophys. J.* **236**, 135-142 (1980).
31. Chevalier, R. A. *Astrophys. J.* **208**, 826-828 (1976).
32. Wenzel, W. *IAU Circ. No.* 4049 (1985).
33. Filippenko, A. V. & Sargent, W. L. W. *Astr. J.* (submitted).

Molecular events during maturation of the immune response to oxazolone

C. Berek, G. M. Griffiths* & C. Milstein

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Sequence analysis of the heavy- and light-chain messenger RNA of hybridomas immunized with a specific hapten yields important clues about the interplay between genetic and selective events during the onset and maturation of the immune response. The maturation of the primary response to the hapten 2-phenyl-5-oxazolone is characterized by a drift to higher-affinity somatic variants of a germline-encoded basic sequence, whereas hybridomas from the secondary response demonstrate a further maturation dominated by a shift to alternative germline combinations.

THERE are good reasons to believe that drastic changes occur during the maturation¹ of an immune response and that primary- and secondary-response antibodies produced to the same antigen are structurally very different. The early stages of a response to an antigenic stimulus are often of a restricted nature²⁻⁴, and followed by a change to an apparently more diverse response correlated with an increase in affinity for the antigen. The primary antibodies to (4-hydroxy-3-nitrophenyl)acetyl (NP) have λ chains and are heteroclitic, whereas the hyperimmune antibodies have κ chains, are no longer heteroclitic and show a high affinity for the immunizing antigen^{5,6}. Similarly, antibodies to phosphorylcholine coupled to keyhole limpet haemocyanin (KLH) from the primary and secondary responses have different structures as well as increased affinity^{7,8}. However, the difference in fine specificity suggests that the structural changes are probably associated with a change in hapten presentation (cell wall of microorganisms and KLH).

More extensive structural information is essential to elucidate the mechanism of known genetic processes that generate antibody diversity (reviewed in ref. 9) in response to a single antigenic stimulant. Antibody diversity arises in several ways: (1) germline diversity resulting from the multiple copies of variable (V), diversity (D) and joining (J) segments; (2) combinatorial diversity resulting from the use of alternative combinations of

these genes; (3) junctional diversity resulting from variations in the joining of the gene segments; (4) somatic point mutations throughout the V regions. How do these mechanisms contribute to the maturation of an immune response?

To study this question, we have isolated hybridomas generated at three stages of the immune response to the hapten 2-phenyl-5-oxazolone (phOx) and determined the sequence of the messenger RNA for the corresponding antibody heavy (H) and light (L) chains. The response at day 7 was found to be restricted. Most of the antibody-secreting cells expressed a H/L chain combination encoded by a pair of germline genes, $V_H\text{-OxI}$ and $V_L\text{-OxI}$, respectively¹⁰. One week later (14 days after the primary injection) most antibodies still expressed similar, but no longer identical, sequences, suggesting a high degree of somatic mutation¹¹. These changes were associated with an increase in affinity for the hapten.

Here, we report the analysis of the secondary response. We were surprised to find that the germline gene combination and variants that characterized the primary response constitute only a minor fraction of the phOx-specific hybridoma lines. The sequences demonstrate that what characterizes the secondary response is the dominant role of new germline H/L-chain gene combinations. We suggest that, although somatic mutation is a major factor in the genetic drift towards better antibodies characteristic of the maturation of the immune response to phOx, the shift towards alternative germline gene expression is equally important in the strategy for obtaining further improvements in antibody affinity at later stages of the response.

* Present address: Department of Pathology, Stanford University Medical Center, Stanford, California 94305, USA.

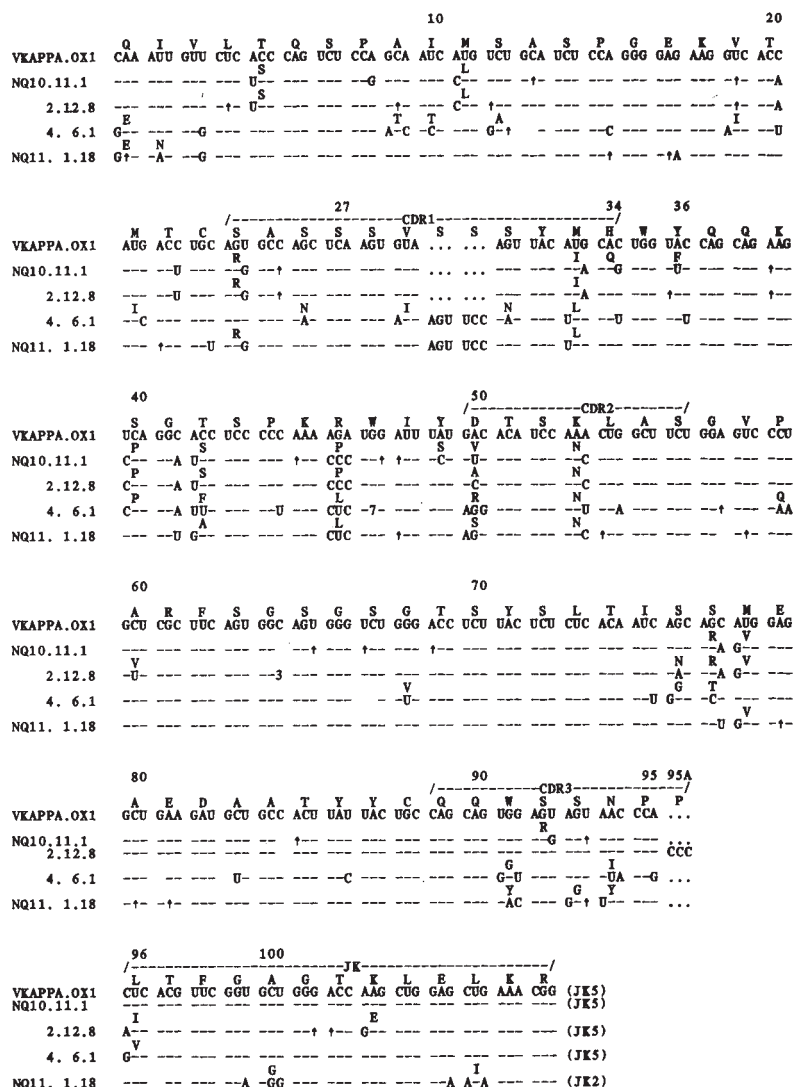


Fig. 1 Immunoglobulin light-chain mRNA sequences for secondary hybridoma lines with specificity for phenylxazolone: V_{κ} -Ox1-related sequences. The top line shows the nucleotide sequence of the germline gene V_{κ} -Ox1 joined to the $J_{\kappa}5$ gene segment (ref. 15 and J. Even, unpublished results). The amino-acid translation is shown above, with numbering according to Kabat²². The complementarity-determining regions (CDR1, 2, 3) and the J segment are indicated. The V_{κ} sequences of the different hybridoma lines are given beneath. Dashes indicate identity with the top line, an arrow marks probable identity, a gap that the nucleotide could not be identified. A computer code is used to denote sequence ambiguities²³: 1, probably C; 2, probably T; 3, probably A; 4, probably G; 5, A or C; 6, G or T; 7, A or T. Established nucleotide differences and the predicted amino-acid changes are indicated.

Methods. Mice were immunized intraperitoneally with 30 µg alum-precipitated pHox coupled to the carrier chicken serum albumin (pHox-CSA)²¹ and 6 weeks (fusion NQ10) or 8 weeks (fusion NQ11) later boosted with an intravenous injection of 100 µg pHox-CSA. Three days later, spleen cells of individual mice were fused with the non-secretor line NSO and hybridoma lines screened for secretion of pHox-specific antibodies. Isolation and sequencing of the mRNA were done as described previously^{11,24}

Secondary response to oxazolone

Hybridomas of the secondary response were prepared from mice boosted 6 weeks (fusion NQ10) or 8 weeks (fusion NQ11) after primary immunization with pHx coupled to chicken serum albumin (pHx-CSA)(see Fig. 1 legend). In both experiments there was a strong secondary response yielding ~30% of hybridomas specific for the hapten pHx. Several of these lines were randomly taken, cloned and the mRNA of the H and the L chain sequenced.

In the secondary response, only 4 out of the 23 lines were related to the V_L - $Ox1$ with V_H - $Ox1$ combination¹¹, indicating a shift to new germline H/L-chain combinations (Table 1). The expression of H chains was more complex than L chains and V_H subgroups emerged which had not been detected at earlier stages of the response (Table 1). For this reason we will analyse the data classified by the criteria of L-chain subgroups.

V_r-Ox subgroup antibodies

Most of the L chains of the V_{κ} -Ox subgroup are derived from a single germline gene, V_{κ} -Ox1^{10,11}. Sequences closely related to V_{κ} -Ox1 but probably derived from other germline genes were also detected throughout the response (Table 1). These are the L chains of NQ2/6.1 and NQ2/45.10.4 from day 7 (ref. 10) and NQ10/11.1, NQ10/2.12.8, NQ10/4.6.1 and NQ11.18 from the secondary response (Fig. 1). These L chains show an 85–95% homology with the Ox1 prototype sequence. We believe that these V_{κ} segments originate from different germline genes. Compared with V_{κ} -Ox1, NQ11/1.18 and NQ10/4.6.1 (originating from different mice) have an identical

insertion of six extra residues in the first complementarity-determining region (CDR1). Furthermore, NQ10/2.12.8 and NQ10/11.1 show the same residue at 15 positions which differ from the prototype *V_κ-Ox1* sequence. Although these two lines originate from the same mouse, they are independent cell lines because the L chains are combined with very different H chains (Fig. 2).

In the secondary response, L chains of the V_{κ} O $\times$ subgroup are often associated with V_H segments which have not been observed in the primary response (Table 1). These V_H sequences differ considerably from those encoded by V_H -Ox1. The homology is 70% at the nucleotide level and only 60% at the translated amino-acid level. Four of them (NQ10/12.5, NQ10/11.1, NQ10/15.3 and NQ11/8.1) belong to the V_H group 5 (MOPC21 gene family)^{12,13}, NQ11/7.12 and NQ11/1.18 to the V_H group 1 (J558 gene family), and NQ10/2.12.8 and NQ10/4.6.1 to the V_H group 3 (germline gene V_H -1210.7)¹⁴ (Fig. 2). V_H regions encoded by the MOPC21 or J558 gene family are one amino acid longer in CDR2 (residue 52A) than V_H -Ox1.

Other V_κ subgroups

Two of the day 7 hybridomas did not express L chains of the V_{κ} -Ox subgroup. The L chain of one of them (NQ2/45.1) was also found in the day 14 hybridoma NQ7/47.1, and in 10 out of 23 of the secondary response hybridomas (Table 1). Most of the sequences appear identical to V_{κ} -45.1 (Fig. 3) except for the pair NQ10/2.22 and NQ10/12.22, which differed at three positions from the V_{κ} -45.1 prototype sequence, as well as NQ10/2.12.4, an IgM hybridoma which has one substitution at the third base of triplet 34. The last base of the V_{κ} -region

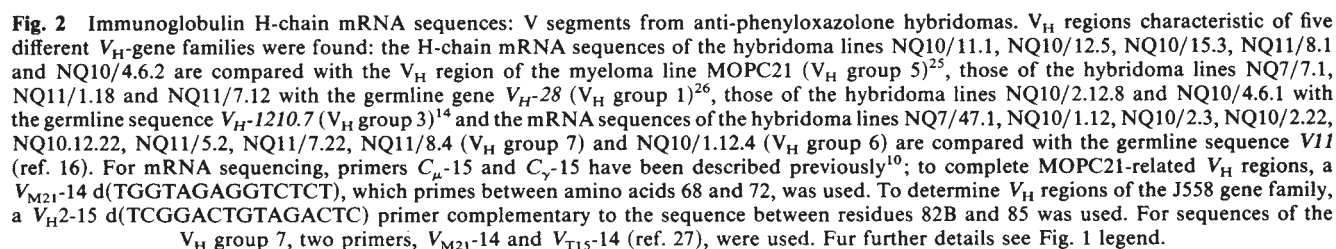


Fig. 4 L-chain mRNA sequences related to V_{κ} -ars. Sequences are compared with the L-chain sequence of the day 7 hybridoma line NQ5/89.4 (ref. 10). For further details see Fig. 1 legend.

Hybridomas expressing a V_{κ} -45.1 L chain again have closely

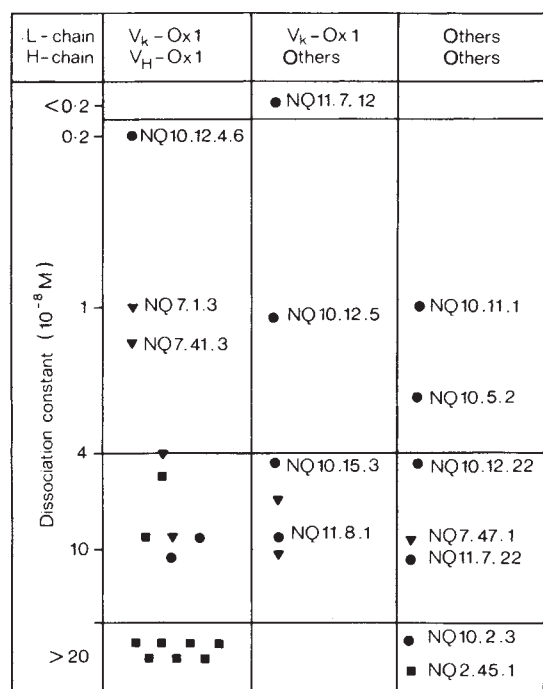


Fig. 6 Dissociation constants of phOx-specific hybridoma proteins. Dissociation constants of hybridoma lines obtained at three different stages of the immune response to phOx (day 7 (■), day 14 (▼), primary and secondary (●) response) are shown. Data are plotted according to the V_H/V_κ combination of the antibody molecules. Symbols with no designation show results from hybridoma lines described previously^{10,11}. Affinity measurements were performed by the method of fluorescence quenching with the hapten phOx-cap²⁹. Antibodies with affinities $<4 \times 10^{-8}$ M were further analysed by the method of equilibrium dialysis, using ³HphOx-aminobutyrate.

related D-J region sequences. In 9 out of 12 the V_H region is followed by D segments 5 amino-acid residues long (Fig. 5 and ref. 9). These include hybridomas of three independent fusions and one (NQ10/2.12.4) expressing a H chain of a different V_H -gene family. All these have a J_{H3} segment of 16 residues. As discussed above, the strong correlation of the length of CDR3/FR4 in H chains of antibodies of a given specificity points to the importance of the interaction between this region and the L chains for the structure of the antibody-combining site.

Affinity maturation

To measure the affinity of the anti-phOx antibodies, we have used the method of fluorescence quenching¹⁷ by 2-phenyl-5-oxazolone aminocaproyl (phOx-cap) (Fig. 6). This method becomes inaccurate for affinities approaching 10^{-8} M. To determine affinities below 4×10^{-8} M more accurately, we used equilibrium dialysis. Antibodies of the early primary response had dissociation constants of $\sim 3 \times 10^{-7}$ M. In only one example was it significantly below 10^{-7} M (ref. 10). During the late primary response almost all of the antibodies had dissociation constants well below 10^{-7} M (ref. 11). In the secondary response, some of the antibodies showed a further increase in affinity. Antibodies NQ10/12.4.6 (V_κ -Ox1- V_H -Ox1) and NQ11/7.12 (V_κ -Ox1- V_H group 1) were at least fivefold better than the two best of the day 14 response. What we find remarkable is that despite the increased number of somatic mutations¹¹, only one out of three secondary-response Ox1 antibodies had affinity constants below 4×10^{-8} M (Fig. 6). On the other hand, three out of four secondary-response antibodies expressing V_κ -Ox1 in combination with a H chain other than V_H -Ox1, had similar high affinity.

Taken together, these results suggest that further point mutations of the V_H -Ox1- V_κ -Ox1 combination did not necessarily lead to an improvement in the affinity of the hapten beyond that achieved at day 14. Other germline combinations were a fresh source for such improvements.

Diversity and response maturation

Three stages of the maturation of the antibody response to oxazolone have been studied. Each stage involved the sequence analysis of the L- and H-chain mRNA of between 11 and 23 hybridomas. In the first stage (7 days after primary immunization) the hybridomas reflect the selection of a germline gene rearrangement which happens to display good affinity. In the second stage (14 days after primary immunization) the antibodies expressing the unmutated germline genes characteristic of day 7 disappear and are largely replaced by drift of the same germline genes, modified by point mutations in all segments of the V regions including the J segments. The point mutations tend to accumulate at certain positions (for example, residues 34 and 36 of V_κ -Ox1) even in hybridomas of different clonal origin, suggesting mutational hotspots and/or a strong selective pressure to accommodate hitherto unknown interactions with ligands, probably antigen itself. Although mutant forms of the V_H -Ox1- V_κ -Ox1 combination constitute most of the late primary-response hybridomas, a shift to alternative germline combinations begins to become apparent. This shift is the dominant feature of the secondary-response hybridomas. At this third stage, only a few hybridomas express the V_H -Ox1- V_κ -Ox1 combination and these show further mutational drift. These additional mutations do not always correlate with an increase in affinity for the hapten. The V_κ -Ox1- V_H -Ox1 combination, although useful to start with, may become increasingly difficult to mutate towards further improvement of binding. On the other hand, other germline gene combinations may be a better starting point from which high-affinity antibodies can be produced following a few key mutations.

The question arises of why these combinations, which are only rarely seen in late primary responses (for example, V_H group 7 and V_κ -45.1), become dominant in the secondary. A calculation can help to clarify the problem. If we assume a total of 200 V genes and 4 J segments for each chain and 20 D gene segments, random assortment results in 800 κ and 16,000 H chains. In addition, there is junctional diversity, which increases by at least a factor of 2.5 the number of possible κ chains to 2,000. As for the H chain, the same factor of 2.5 for junctional diversity is applicable to V/D and D/J boundaries. Considering variations in the length for both J and D segments¹⁰, an additional factor of 5 is an underestimate. Without bringing into the calculation the potential diversity generated by the postulated N regions¹⁸, this takes the total of potential H-chain V regions, quite apart from somatic variants, to 500,000. The combinatorial association of κ and H chains gives a total of 10^9 , more than the total number of lymphocytes in a mouse. Although the combinatorial expression is unlikely to be random¹⁹ and not all junctional or combinatorial diversity will be structurally viable, it is possible that the mouse does not express at any given moment all the variants which it can potentially generate by using the unmutated germline genes. It follows that the first expression of the response will be dominated by the frequency at which certain combinations occur as much as by the affinity for the antigen. The expression of a low-frequency germline gene combination of an infrequent joining variant, or even a random favourable point mutation, could eventually arise and start the shift to a new germline set. Persistence of antigen expands these rare cells into a proliferating population from which mutants can be selected. A shift to new germline gene combinations may have as a secondary effect an increase in potential cross-reactive antibodies. This could explain the apparent paradox of an increase in the cross-reactivity of the antiserum of a secondary response²⁰ while the monoclonal antibodies derived from it have increased affinity (that is, specificity) for the immunogen.

In summary, we suggest that the initial selection of germline gene combinations is based both on frequency and affinity. This is followed by two events. One, selection of somatic mutants of the initial proliferating dominant clones expressing those genes, and second, the active recruitment of other germline gene

combinations. Among these new germline gene combinations, those having the best potential to be modified by somatic mutation into more suitable combining sites will give rise to the very high-affinity antibodies that characterize the mature response. Our experiments do not allow us to say whether the shift to other germline genes is selected by antigen alone or whether there is some other mechanism which favours that shift, such as anti-idiotypic suppression of the initial idotype,

separate compartmentalization of memory and virgin B cells or senescence of early-stimulated clones.

We thank Mr J. Jarvis for measuring the affinity of the antibody molecules, Dr M. Wayne for preparation of oligonucleotide primers, Dr R. Staden for the computer analysis, Dr M. Kaartinen for gifts of oxazolone derivatives, Dr M. Neuberger for helpful discussions, and Miss J. Firth for preparation of the manuscript.

Received 18 March; accepted 13 June 1985.

- Siskind, G. W. & Benaceraff, B. *Adv. Immun.* **10**, 1-50 (1969).
- Cosenza, H. & Köhler, H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2701-2705 (1972).
- Nisonoff, A., Ju, S. T. & Owen, F. L. *Immun. Rev.* **34**, 89-118 (1977).
- Jack, R. S., Imanishi-Kari, T. & Rajewsky, K. *Eur. J. Immun.* **7**, 559-565 (1977).
- Mäkelä, O. & Karjalainen, K. *Transplant. Rev.* **34**, 119-138 (1977).
- Reth, M., Hämmerling, G. J. & Rajewsky, K. *Eur. J. Immun.* **8**, 393-400 (1978).
- Chang, S. P., Brown, M. K. & Rittenberg, M. B. *J. Immun.* **128**, 702-706 (1982).
- Chang, S. P. *et al. J. Immun.* **132**, 1550-1555 (1984).
- Tonegawa, S. *Nature* **302**, 575-581 (1983).
- Kaartinen, M., Griffiths, G. M., Markham, A. F. & Milstein, C. *Nature* **304**, 320-324 (1983).
- Griffiths, G. M., Berek, C., Kaartinen, M. & Milstein, C. *Nature* **312**, 271-275 (1984).
- Dildrop, R. *Immun. Today* **5**, 100-101 (1984).
- Brödeur, P. H. & Riblet, R. *Eur. J. Immun.* **14**, 922-930 (1984).
- Near, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2167-2171 (1984).
- Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-299 (1979).
- Crews, S., Griffin, J., Huang, H., Calame, K. & Hood, L. *Cell* **25**, 59-66 (1981).
- Parker, C. W. *Handbook of Experimental Immunology* 3rd edn, Ch. 18 (Blackwell Scientific, Oxford, 1978).
- Alt, F. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
- Yancopoulos, G. D. *et al. Nature* **311**, 727-733 (1984).
- Sperling, R., Francus, T. & Siskind, G. W. *J. Immun.* **131**, 882-885 (1983).
- Kaartinen, M. *et al. J. Immun.* **130**, 937-945 (1983).
- Kabat, E. A. *et al. Sequences of Proteins of Immunological Interest* (NIH, Bethesda, 1983).
- Staden, R. *Nucleic Acids Res.* **10**, 4731 (1982).
- Griffiths, G. M. & Milstein, C. in *Hybridoma Technology in the Biosciences and Medicine* (ed. Springer, T. A.) (Plenum, New York, in the press).
- Bothwell, A. L. M. *et al. Cell* **23**, 625-637 (1981).
- Loh, D. Y., Bothwell, A. L. M., White-Scharf, M. E., Imanishi-Kari, T. & Baltimore, D. *Cell* **33**, 85-93 (1983).
- Berek, C. *Eur. J. Immun.* **14**, 1043-1048 (1984).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
- Makela, O., Kaartinen, M., Pelkonen, J. L. & Karjalainen, K. *J. exp. Med.* **148**, 1644-1660 (1978).

LETTERS TO NATURE

Anatomy of a cosmic-ray neutrino source and the Cygnus X-3 system

F. W. Stecker, A. K. Harding & J. J. Barnard

Laboratory for High Energy Astrophysics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

There is strong evidence that a compact object in the Cygnus X-3 binary system produces an intense beam of ultra-high-energy cosmic rays. Here, we examine the effects of such a beam hitting the companion star and of the subsequent production of secondary neutrinos. We consider how high a beam luminosity is allowed and how high a neutrino to γ -ray (ν/γ) ratio can be obtained from such a system. We find a maximum allowable beam luminosity of $\sim 10^{42}$ erg s $^{-1}$ for a system consisting of a compact object and a $\sim 1-10 M_{\odot}$ main-sequence target star. The proton beam must heat a relatively small area of the target star to satisfy observational constraints on the resulting stellar wind. With such a model, a ν/γ flux ratio of $\sim 10^3$ can result from a combination of γ -ray absorption and a large ν/γ duty cycle ratio. We find that the high density of the atmosphere resulting from compression by the beam leads to pion cascading and a neutrino spectrum peaking at 1-10 GeV energies, which may avoid catastrophic heating of the target star through internal ν interactions. The ν flux and duty cycle are predicted to be accordingly reduced in the energy range above 1 TeV available to a deep underwater neutrino detector.

There has recently been much interest in constructing theoretical models for the Cyg X-3 binary system¹⁻³ in view of the discovery of ultra-high-energy γ rays from this source⁴⁻⁶. The γ -ray flux implies that Cyg X-3 is the first identified source of ultra-high-energy cosmic rays and that the source power is at least 10^{39} erg s $^{-1}$ in cosmic-ray primaries, enough to provide a significant fraction of the 10^{17} -eV cosmic rays in the Galaxy⁷. It has been suggested that Cyg X-3 may be a source of cosmic-ray neutrinos^{8,9}, produced from the pion decay processes¹⁰ which probably produce the γ rays. This possibility has taken on more interest with the detection of muon signals in deep underground proton decay detectors (ref. 11 and J. Learned, personal communication) from the direction of Cyg X-3 with the 4.8-h periodicity of the Cyg X-3 system. (Such detectors make use of the Cerenkov light emitted by charged particles crossing large volumes of water¹².) The very detectability of such signals, if induced by cosmic-ray neutrinos from Cyg X-3, implies a sur-

prisingly large neutrino flux, primary cosmic-ray beam power ($\gg 10^{39}$ erg s $^{-1}$) and neutrino-to- γ -ray flux ratio. In the event that the recently detected muons are produced by some other 'X' particle, a large beam power and accompanying neutrino flux are almost certainly implied. Therefore, regardless of the outcome of the present observational situation, it is of interest to consider the theoretical implications of significantly larger beam power.

Hillas² has recently presented calculations of γ -ray production from an energetic proton beam, based on a model originally proposed by Vestrand and Eichler¹³. In this model, a beam of $\sim 10^{17}$ eV protons, accelerated at a compact object, impinges on the atmosphere of a companion star, initiating a cascade producing γ rays below 10^{16} eV. The total luminosity of the compact source required to account for the observed γ rays above 10^{12} eV is $L_{\gamma} = 6.4 \times 10^{39} (\epsilon_{\gamma}/0.1)^{-1} (\langle \Omega_{\beta} \rangle / 4\pi) (0.025/\Delta_{\gamma})$ erg s $^{-1}$, where $\langle \Omega_{\beta} \rangle$ is the time-averaged solid angle of the proton beam, ϵ_{γ} is the fraction of proton energy converted into γ rays and Δ_{γ} is the duty cycle of the pulse ~ 0.025 (ref. 5). Owing to observational uncertainty, a smaller Δ_{γ} cannot be ruled out, which would imply an even larger beam luminosity. If the part of the beam luminosity striking the target star, L_{β} , is greater than its Eddington limit,

$$L_{\text{edd}} = \frac{4\pi GcM_{*}}{\kappa} = 1.3 \times 10^{38} \tilde{M} \text{ erg s}^{-1} \quad (1)$$

where M_{*} is the target star mass and κ is the electron scattering opacity, a mass loss will result. (Note that all quantities with overbars will be in solar units, for example, $\tilde{R} = R/R_{\odot}$.) The mass loss from the system is limited by the observed derivative, $\dot{P}$, of the 4.8-h orbital period¹⁴, through the relation¹⁵

$$\dot{P}/P = 2\dot{M}/M_T \quad (2)$$

where M_T is the total mass of the system and $\dot{M}$ is the mass loss rate from the target star (assuming that the orbital angular momentum is lost from the system and that the orbit decay from gravitational radiation is small). When the beam hits the atmosphere of the target star, roughly half of the beam energy flux will be thermalized in the stellar atmosphere, producing an outward radiative flux. The resulting radiation pressure can drive a wind with a terminal velocity v given by

$$v^2 = \frac{2}{R_{*}} \left(\frac{L_{\beta}\kappa}{4\pi fc} - GM_{*} \right) \quad (3)$$

where R_{*} is the target star radius and f is the fraction of the

area of the star illuminated by the proton beam. Energy conservation requires that

$$L_B \approx L_{\text{rad}} + \frac{GM_* \dot{M}}{R_*} + \frac{3}{2} \dot{M} \left(\frac{kT}{m} \right) + \frac{1}{2} \dot{M} v^2 \quad (4)$$

where m is the mean atomic mass of the atmosphere. The radiative flux, L_{rad} , can be neglected if $L_B \gg L_{\text{edd}}$, as can the thermal energy flux (third term). The mass loss rate, combining equations (3) and (4), is then

$$\dot{M} = 4\pi f R_* c / \kappa = 9.6 \times 10^{-4} f \tilde{R} M_\odot \text{ yr}^{-1} \quad (5)$$

The mass loss rate is independent of the beam luminosity and the mass of the target star (provided that $L_B > L_{\text{edd}}$). Using equations (2) and (5), we can derive a relation between R_* and M_T : $R_* = 1.5 \times 10^8 f^{-1} \tilde{M}_T \text{ cm}$. This relation is valid only in the case where the wind from the target star is lost from the system. In this case, the escape velocity of the system is less than v and $f < 12 \tilde{R}^{-1} \tilde{M}_T^{-3/2} L_{39}$. Therefore, since $f < 1$ and $L_B > 10^{39} \text{ erg s}^{-1}$, nearly all of the wind escapes the system. Equations (2) and (5) may then be used to find the fraction of the area of the star heated by the beam such that the resulting mass loss gives the observed $\dot{P} = 1.2 \times 10^{-9}$ (ref. 14):

$$f = 2.2 \times 10^{-3} \tilde{M}_T \tilde{R}^{-1} \quad (6)$$

The angular extent θ of the time-averaged beam in one direction is given by $\tan \theta = 2\pi f / (a/R_* - 1)$. Since a/R_* may be very close to unity (as indicated from the X-ray light curve), the small area predicted by equation (6) does not necessarily imply a small beam angle. The X-ray light curve shows deep but not sharp eclipses^{16,17}, thus the optical depth of the outflowing material to electron scattering is $\tau_e \sim 1$ at distances of the order of the orbital separation, $a = 10^{11} \tilde{M}_T^{1/3} \text{ cm}$. The optical depth τ_x at a distance r is $\tau_x = \kappa \rho r$, where ρ is the mass density of the wind, assuming that v is constant for $r \gg R_*$. Then using equation (6), and mass conservation in the wind, $\dot{M} = 4\pi f r^2 \rho v$, the radius at which $\tau_x = 1$ is $R_1 = (c/v) R_*$ and

$$\frac{R_1}{a} = 5.3 \tilde{M}_T^{1/6} \tilde{R} L_{39}^{-1/2} \quad (7)$$

Thus, the condition $R_1/a \sim 1$ required by the X-ray light curve and suggested by cocoon models for Cyg X-3 (refs 16, 17) can be satisfied.

When the beam grazes the stellar limb, it can temporarily heat an area of $\sim 4\pi \tilde{R}_*^2$ to a temperature of

$$T = 6 \times 10^5 \tilde{M}_T^{-1/4} \tilde{R}^{-3/4} L_{39}^{1/4} \text{ K} \quad (8)$$

This increases the atmospheric scale height to

$$h = kT/mg \approx 2 \times 10^9 L_{39}^{1/4} \tilde{M}_T^{-1/4} \tilde{R}^{5/4} \tilde{M}^{-1} \text{ cm} \quad (9)$$

where g is the surface acceleration due to gravity. This is a lower limit on h , as radiation pressure will tend to increase the scale height. The primary protons have a mean free path for interaction of $\Lambda_p \sim 30 \text{ g cm}^{-2}$ at high energies¹⁸. They produce π^0 mesons which decay into γ rays having a mean free path against pair production off electrons in the stellar atmosphere of $\sim 50 \text{ g cm}^{-2}$ at an energy of $\sim 1 \text{ TeV}$. Because of absorption effects, the γ rays are strongly produced only at the limb of an atmosphere¹⁹. The primaries also produce charged pions which decay into neutrinos. The efficiency for neutrino production reaches a plateau⁹ at an atmospheric depth of $5\Lambda_p$.

Whereas in the case where the beam is at the stellar limb, heating of the atmosphere increases the scale height h , in the case where the beam hits the star directly, ram pressure from the beam with energy flux F_B , acting downward in concert with gravity, decreases h . In this case, h is given by

$$h = \frac{kT}{m} \left[\frac{F_B}{5\Lambda_p c} \delta + \frac{GM_*}{R_*^2} \right]^{-1} \approx 3 \times 10^7 L_{39}^{-3/4} \tilde{R}^{1/4} \tilde{M}_T^{3/4} \text{ cm} \quad (10)$$

where we have assumed energy and momentum balance in the

atmosphere. The parameter $\delta \approx (1 - \Phi_u/\Phi_T)$, where Φ_T is the total radiative flux and Φ_u is the upward flux component, takes values between 0 and 1, depending on geometry and radiation transfer in the atmosphere. When $\delta = 0$, all of the radiation propagates upwards, producing a force which directly opposes the ram force, and no compression occurs. When $\delta = 1$, all radiation propagates or escapes sideways and the atmosphere will be compressed by the full ram force of the beam. However, at the edges of the beam, the escaping radiation should drive the mass loss estimated by equation (5). Since the smallest dimension of the beam, $x = 2\pi f R_* \approx 9 \times 10^8 \text{ cm}$, is less than the uncompressed atmospheric scale height as derived in equation (9), we can assume that $\delta \sim 1$. Because the beam encounters much higher grammage here, the γ rays are completely absorbed. For ν energies below $\sim 100 \text{ TeV}$, the cross-section for interactions with nucleons is in the linear region¹⁰ and is given by (expressing all energies in TeV units) $\sigma_\nu = 7 \times 10^{-36} E_\nu \text{ cm}^2$. The optical depth of a star to neutrinos is then given by $\tau_\nu = 0.41 E_\nu \tilde{M} \tilde{R}^{-2}$, so that the critical energy above which the star is optically thick to neutrinos is $E_\nu^* = 2.4 \tilde{M}^{-1} \tilde{R}^2 \text{ TeV}$.

If neutrinos are generated above E_ν^* , such neutrinos will be absorbed within the main body of the star uniformly, converting into high-energy muons and electrons which heat up the inside of the star where the photon diffusion escape time is $\sim 10^6 \text{ yr}$. Absorption of an energy flux significantly above the Eddington limit would result in disruption of the target star on a timescale $< 10^3 \text{ yr}$. Therefore, such neutrinos must not be generated in a plausible theoretical scenario.

Fortunately, ν production above $\sim 1 \text{ TeV}$ energy can be suppressed. Owing to relativistic time dilation, the distance travelled by the parent pion before decay as a function of energy is given by $l_\pi = 5.57 \times 10^9 E_\pi \text{ cm}$. If this distance is greater than the interaction mean free path, the pion will interact and produce a pion cascade rather than decay into neutrinos directly²⁰. This will result in the suppression of higher-energy neutrinos and the multiple production of neutrinos of low enough energy (from the decay of lower-energy parent pions) that they can travel through the star without being absorbed. This situation will occur if the atmospheric density at a depth of $\sim 50\text{--}200 \text{ g cm}^{-2}$ (roughly where ν production is most efficient) is greater than $(l_\pi \sigma_\pi)^{-1}$ where $\sigma_\pi (\geq 1 \text{ TeV}) \approx 6 \times 10^{-26} \text{ cm}^2$ (ref. 18), corresponding to a critical energy, $E_\nu^c \approx (10^{-6} \text{ g cm}^{-3}/\rho) \text{ TeV}$. For $E_\nu < E_\nu^c$, we find the required density criterion for the neutrino production region should be $\rho < \rho_c = 4.2 \times 10^{-7} \text{ g cm}^{-3} \tilde{M} \tilde{R}^{-2}$. From the scale height in equation (10), we have $\rho \approx 5\Lambda_p/h = 5 \times 10^{-6} L_{39}^{3/4} \tilde{M}_T^{-3/4}$. This effectively suppresses the more dangerous ν s since $E_\nu^c < E_\nu^*$. Because the energy threshold for ν detection with DUMAND²¹ is $\sim 1 \text{ TeV}$, we predict that Cyg X-3 may not be a superstrong source for such a detector.

The fraction of π -decay ν s that interact even when $\tau_\nu \ll 1$ will contribute to heating the stellar interior. The amount of energy absorbed by the star is $\epsilon_\nu \tau_\nu L_B \sim 4.1 \times 10^{38} \epsilon_\nu L_{39} \tilde{M} \tilde{R}^{-2} (E_\nu) \text{ erg s}^{-1}$, where ϵ_ν is the efficiency for neutrino production. The average energy $\langle E_\nu \rangle \approx 1\text{--}10 \text{ GeV}$ for $\rho = 10^{-6} \text{ g cm}^{-3}$ (ref. 20). In this case, detailed cascade calculations required to obtain a reliable heating estimate may impose severe constraints on the Hillas model.

There is a limit to high-energy ν suppression. There will always be high-energy ν s from prompt decay of mesons carrying charm or higher mass quark flavours. The ratio ξ of forward production of these mesons to π s is probably $\sim 10^{-4}\text{--}10^{-3}$ at high energies, although the experimental situation is still in some doubt²². Taking $\xi \geq 10^{-4}$, we can use the existence of prompt neutrinos to place a theoretical upper limit on the beam power, L_B .

Let us assume that the π -decay neutrinos are effectively suppressed at high energies by ram-pressure density enhancement in the region of the stellar atmosphere directly below the particle beam where the π s are produced (see equation (10)). We are then still left with the prompt ν s which will give a lower limit to the high-energy ν flux. As about half the beam energy goes into first generation ν s from π s (or their cascade

equivalent), taking $\xi \geq 10^{-4}$ implies that at least $\sim 5 \times 10^{-5} L_B$ ($> E_p^*$) will go into prompt ν s that heat the stellar interior, where $E_p^* \approx 4E_p^*$ and E_p^* is the energy above which σ_p is large enough to assure $\tau \geq 1$. For main-sequence stars of $1M_\odot$ and $10M_\odot$, $E_p^* \sim 10$ and ~ 25 TeV respectively. In the Hillas² model, essentially all of the beam power lies in this energy range, so that the interior heating rate from prompt ν s is given by $\eta_{pr} \geq 5 \times 10^{-5} L_B$. If $\eta_{pr} > L_{edd} \approx 10^{38} M \text{ erg s}^{-1}$, the target star would be disrupted. This places an upper limit on the beam power for the Cyg X-3 system:

$$L_B \leq 2 \times 10^{42} \tilde{M} \text{ erg s}^{-1} \quad (11)$$

Unless $\eta_{pr} < L_*$, the intrinsic stellar luminosity, the structure of the star will be modified and its radius will expand. For main-sequence stars, this condition is given by

$$L_B \leq \begin{cases} 8 \times 10^{37} \text{ erg s}^{-1}, & \tilde{M} = 1 \\ 2 \times 10^{40} \text{ erg s}^{-1}, & \tilde{M} = 4 \\ 4 \times 10^{41} \text{ erg s}^{-1}, & \tilde{M} = 10 \end{cases} \quad (12)$$

This can be compared with the π -decay neutrino heating limit, $\eta_\pi \equiv \epsilon_\nu \tau_\nu L_B < L_*$,

$$L_B \leq \begin{cases} 10^{38} (\epsilon_\nu/0.1)^{-1} (E_\nu/10^{-3})^{-1} \text{ erg s}^{-1}, & \tilde{M} = 1 \\ 4.7 \times 10^{40} (\epsilon_\nu/0.1)^{-1} (E_\nu/10^{-3})^{-1} \text{ erg s}^{-1}, & \tilde{M} = 4 \\ 1.3 \times 10^{42} (\epsilon_\nu/0.1)^{-1} (E_\nu/10^{-3})^{-1} \text{ erg s}^{-1}, & \tilde{M} = 10 \end{cases} \quad (13)$$

Should either prompt neutrino or π -decay heating exceed the stellar luminosity, radiative transfer will no longer be efficient enough to carry the increased luminosity to the surface and the star will convectively transport this additional energy. The star will expand (see ref. 23), returning to the Hayashi track in the Hertzsprung-Russell diagram, until cooling to a surface temperature of ~ 3 to 4×10^3 K. The new radius which results is $R_* \sim 7 \times 10^{13} \eta_{39}^{1/2} (T/4,000 \text{ K})^{-2} \text{ cm}$. This R_* would then exceed the stellar orbit, and may quench the beam of the compact object, reducing the radius until it is within the size of the orbit. The stability of such a system is unknown.

We now consider the ratio of ν -to- γ -ray fluxes in the case where the above constraints on L_B are satisfied. There are two situations:

(1) In the case where the beam is grazing the target star's atmosphere at the limb, for pathlengths $X \leq 25 \text{ g cm}^{-2}$, comparable fluxes of ν and γ rays are produced¹⁰. For astrophysically significant production pathlengths $\sim 60 \text{ g cm}^{-2}$, the ν/γ ratio is considerably higher owing to attenuation of the γ rays by pair production. This flux ratio is approximately an order of magnitude for a cosmic-ray primary spectrum of index near 2 (ref. 9). The fraction of the orbital period during which the stellar limb eclipses the source of the beam will be denoted as Δ_* .

(2) Where the beam is hitting the star directly, the neutrinos produced by the decay of the pions within $X \approx 5\Lambda_p$, pass through the star and exit provided they are below some critical energy E_p^* . Denoting the fraction of the period in which the beam is directly hitting the star by Δ_* , the ratio of neutrino flux to γ -ray flux per cycle is given by $F_\nu/F_\gamma \approx 10\Delta_*/\Delta_a$, where the factor of 10 is due to the ratio of ν -to- γ -ray production efficiencies. The ratio, Δ_*/Δ_a is given by $\Delta_*/\Delta_a \approx R_*/h \approx 100$, with $R_* \approx 10^{11} \text{ cm}$ and $H \approx 10^9 \text{ cm}$ (see equation (9)). Therefore, the ratio of neutrinos to γ rays per orbit can be as high as $\sim 10^3$. Thus, one can obtain much higher ν/γ ratios than previously thought.

Our results therefore indicate that high ν/γ ratios and neutrino fluxes are possible from a source such as Cyg X-3, but that there are also strict theoretical constraints from such models. Such constraints are independent of the details of the models involved and should be useful in considering future observational data.

Note that since submission of this letter two papers^{24,25} on neutrino production in Cyg X-3 have come to our attention.

We thank T. Gaisser, F. Halzen, D. Kazanas, J. Learned, M. Marshak and R. Ramaty for helpful discussions. J.J.B. is a

National Academy of Sciences National Research Council Post-doctoral Resident Research Associate.

Received 11 April; accepted 12 June 1985.

1. Eichler, D. & Vestrand, W. T. *Nature* **307**, 613-614 (1984).
2. Hillas, A. M. *Nature* **312**, 50-51 (1984).
3. Channugam, G. & Brecher, K. *Nature* **313**, 767-768 (1985).
4. Samorski, M. & Stamm, W. *Astrophys. J. Lett.* **268**, L17-L21 (1983).
5. Lloyd-Evans, J. et al. *Nature* **305**, 784-787 (1983).
6. Douthwaite, J. C. et al. *Astr. Astrophys.* **126**, 1-6 (1983).
7. Wdowczyk, J. & Wolfendale, A. W. *Nature* **305**, 609-610 (1983).
8. Eichler, D. *Nature* **275**, 725-726 (1978).
9. Stenger, V. J. *Astrophys. J.* **286**, 810-816 (1984).
10. Stecker, F. W. *Astrophys. J.* **228**, 919-927 (1979).
11. Marshak, M. L. et al. *Phys. Rev. Lett.* **54**, 2079-2082 (1985).
12. Perkins, D. H. A. *Rev. nucl. Part. Sci.* **34**, 1-52 (1984).
13. Vestrand, W. T. & Eichler, D. *Astrophys. J.* **261**, 251-258 (1982).
14. Bonnet-Bidaud, J. M. & van der Klis, M. *Astr. Astrophys.* **101**, 299-304 (1981).
15. Davidsen, A. & Ostriker, J. P. *Astrophys. J.* **189**, 331-338 (1974).
16. Ghosh, P., Elsner, R. F., Weisskopf, M. C. & Sutherland, P. G. *Astrophys. J.* **251**, 230-245 (1981).
17. White, N. E. & Holt, S. S. *Astrophys. J.* **257**, 318-337 (1982).
18. Baltrusaitis, et al. *Phys. Rev. Lett.* **52**, 1380-1383 (1984).
19. Stecker, F. W. *Nature phys. Sci.* **242**, 59-60 (1973).
20. Gaisser, T. K. & Stanev, T. *Phys. Rev. Lett.* **54**, 2265-2268 (1985).
21. Roberts, A. (ed) *Proc. 1978 DUMAND Sum. Wkshp* (Scripps Institution Press, 1979).
22. Kernan, A. & Van Dalen, G. *Phys. Rep.* **106**, 297-398 (1984).
23. Mathews, W. G. *Astrophys. J.* **272**, 390-399 (1983).
24. Lee, H. & Bludman, S. A. *Astrophys. J.* **290**, 28-32 (1985).
25. Kolb, E. W., Turner, M. S. & Walker, T. P. Preprint, Fermi Laboratory (1985).

Nuclear power stations as a background source for antineutrino astronomy

P. O. Lagage

Centre d'Etudes Nucléaires de Saclay, Service d'Astrophysique, 91191 Gif-sur-Yvette Cédex, France

The study of the electronic antineutrinos ($\bar{\nu}_e$) of low energy (≤ 10 MeV) is an active and multidisciplinary field. The first generation of large underground detectors, awaiting bursts of $\bar{\nu}_e$ following a supernova explosion in our Galaxy, is now operating^{1,2}. When placed near nuclear reactors, $\bar{\nu}_e$ detectors can be used to study neutrino oscillations^{3,4}. Geophysicists are also concerned because the Earth is continuously emitting $\bar{\nu}_e$ (ref. 5). Here, I calculate the $\bar{\nu}_e$ flux due to all the nuclear power stations and show that this background must be taken into consideration when studying the detection of low-energy $\bar{\nu}_e$.

A unique feature of the explosion of a supernova is a burst of neutrinos and antineutrinos of energy ~ 10 MeV; $\sim 2 \times 10^{11} \bar{\nu}_e \text{ cm}^{-2}$ are expected on Earth from one supernova exploding in the centre of the galaxy². All the supernova explosions since the beginning of the Universe have formed a faint diffuse $\bar{\nu}_e$ background of $\sim 50 (\bar{\nu}_e \text{ of energy } \sim 6 \text{ MeV}) \text{ cm}^{-2} \text{ s}^{-1}$ (ref. 5); these $\bar{\nu}_e$ provide information on the supernova rate averaged over time and galaxies. Several underground $\bar{\nu}_e$ detectors of 100 tonnes, based on the reactions: $\bar{\nu}_e + p \rightarrow n + e^+$ and $n + p \rightarrow d + \gamma$, are waiting for one supernova to explode in the Galaxy. A 1-ktonne detector is planned in the Gran Sasso laboratory.

As the energy threshold of the reaction $\bar{\nu}_e + p \rightarrow n + e^+$ is 1.8 MeV, this type of detector could also be sensitive to the high-energy tail of the 'natural' $\bar{\nu}_e$ due to the decay of isotopes inside the Earth. Indeed, their maximum energy is 3.3 MeV and the total flux has been estimated⁵ to be $\sim 10^7 \bar{\nu}_e \text{ cm}^{-2} \text{ s}^{-1}$. With an $\bar{\nu}_e$ energy threshold of 2.5 and 3 MeV respectively, a 1-ktonne detector with an efficiency of 70% should count, respectively 15 and four events yr^{-1} which, at these energies, originate only from the ²¹⁴Bi decay.

A potential antineutrino background source for all the previous studies is the nuclear power stations which emit $\bar{\nu}_e$ with energies up to ~ 10 MeV. For example, a flux of $1.5 \times 10^{12} (\bar{\nu}_e \text{ of energy } \geq 3.3 \text{ MeV}) \text{ cm}^{-2} \text{ s}^{-1}$ has been measured at a distance of 13.63 m from a typical power station, the Bugey reactor in France⁴; the corresponding flux in the whole energy range is $2 \times 10^{13} \bar{\nu}_e \text{ cm}^{-2} \text{ s}^{-1}$. The power delivered in 1983 by the 317 nuclear power stations throughout the world was 191 GWe, which is equivalent to the power of 212 Bugey reactors; this

Table 1 Antineutrino fluxes due to the nuclear power stations in different underground laboratories

Location of laboratory	Fluxes of $\bar{\nu}_e$ ($\text{cm}^{-2} \text{s}^{-1}$)	
	Present	Including reactors under construction
Fairport Harbor (Ohio, USA)*	1.3×10^6	$\sim 4.7 \times 10^7$
Mont Blanc tunnel (France)*†	3.0×10^6	4.7×10^6
Frejus tunnel (France)*	3.0×10^6	4.6×10^6
Kamioka (Japan)*	1.6×10^6	2.4×10^6
Artemosk (USSR)†	3.3×10^5	1.1×10^6
Gran Sasso tunnel (Italy)*†‡	4.5×10^5	1.0×10^6
Soudan (Minnesota, USA)*	3.0×10^5	4.2×10^5
Baksan (USSR)†	1.2×10^5	3.4×10^5
Lead (Dakota, USA)‡	1.4×10^5	2.1×10^5
Park City (Utah, USA)*	1.0×10^5	1.7×10^5
Kolar Gold Field (India)*	3.9×10^4	7.3×10^4

*Experiments on the proton lifetime. † Experiments on supernova antineutrinos. ‡ Experiments on solar neutrinos. (The experiments at Gran Sasso are in project only.)

power should double by 1990 (ref. 6). A lower limit to the reactor $\bar{\nu}_e$ background flux is obtained by assuming that the reactors are all placed at the same point on Earth and that the detector is placed at the antipodal point. The total flux anticipated is then $\geq 5 \times 10^3 \bar{\nu}_e \text{ cm}^{-2} \text{ s}^{-1}$ and the flux of $\bar{\nu}_e$ of energy $\geq 5 \text{ MeV}$ is still $\geq 60 \bar{\nu}_e \text{ cm}^{-2} \text{ s}^{-1}$.

This lower limit is of the same order of magnitude as the diffuse $\bar{\nu}_e$ flux expected from all the supernova explosions which have ever occurred. So, the $\bar{\nu}_e$ background from nuclear power stations will make the detection of the diffuse supernova $\bar{\nu}_e$ flux even more difficult. This point is supported by calculating the flux of nuclear-reactor $\bar{\nu}_e$ in the present underground laboratories. As can be seen from Table 1, the reactor $\bar{\nu}_e$ fluxes in these laboratories are at least an order of magnitude higher than the lower limit calculated previously and will be, around 1987, as high as $\sim 10^4$ times this limit at the location of the IMB (Irvine, Michigan, Brookhaven) proton-decay experiment (Fairport Harbor). Thus, the nuclear reactors introduce a significant $\bar{\nu}_e$ background, hence to detect the diffuse supernova $\bar{\nu}_e$, the detectors should have some angular resolution or an energy threshold of $\sim 9 \text{ MeV}$ (see Fig. 1).

Table 1 also shows that the nuclear power stations can be an important source of background for the detection of the Earth's $\bar{\nu}_e$. Indeed, in all the present underground laboratories, the reactor $\bar{\nu}_e$ flux is $\geq 1\%$ of the Earth's $\bar{\nu}_e$ flux; in most cases, it is $\geq 10\%$ and even $> 100\%$ in one case. Note that if we consider only the ^{214}Bi -decay $\bar{\nu}_e$, the importance of the reactor $\bar{\nu}_e$ background is increased by a factor of ~ 50 . Thus, the $\bar{\nu}_e$ background due to the nuclear power stations should be taken into account in siting future low-energy $\bar{\nu}_e$ detectors. Sites with very low reactor $\bar{\nu}_e$ background ($\sim 8 \times 10^3 \bar{\nu}_e \text{ cm}^{-2} \text{ s}^{-1}$ actually, $\sim 1.3 \times 10^4 \bar{\nu}_e \text{ cm}^{-2} \text{ s}^{-1}$ in the near future) could be found, for example, in Australia or New Zealand.

The nuclear reactor $\bar{\nu}_e$ background is important for the future detection of the diffuse supernova $\bar{\nu}_e$ and of the Earth's $\bar{\nu}_e$. How does it compete with the non-antineutrino background in the present and future supernova burst detectors? In the first generation of detectors, the non-antineutrino background is important. For example, in the Mont Blanc detector, the energy threshold for the $\bar{\nu}_e$ has been set to 7 MeV and the mean background is $\sim 20 \text{ counts yr}^{-1} \text{ tonne}^{-1}$ in the inner counters; it is much lower in the core of the detector (ref. 1 and Ryazhskaya UP85 Conf., St Vincent, 1985). This indicates that a better shielding of the detector is possible. The nuclear-reactor $\bar{\nu}_e$ background in the detector represents, at present, less than 1% of the non-antineutrino background. Thus, the reactor $\bar{\nu}_e$ background does not perturb the detection of the $\bar{\nu}_e$ burst following one supernova explosion (see Fig. 1). Future detectors will be larger (several kilotonnes) and better shielded so that the non-

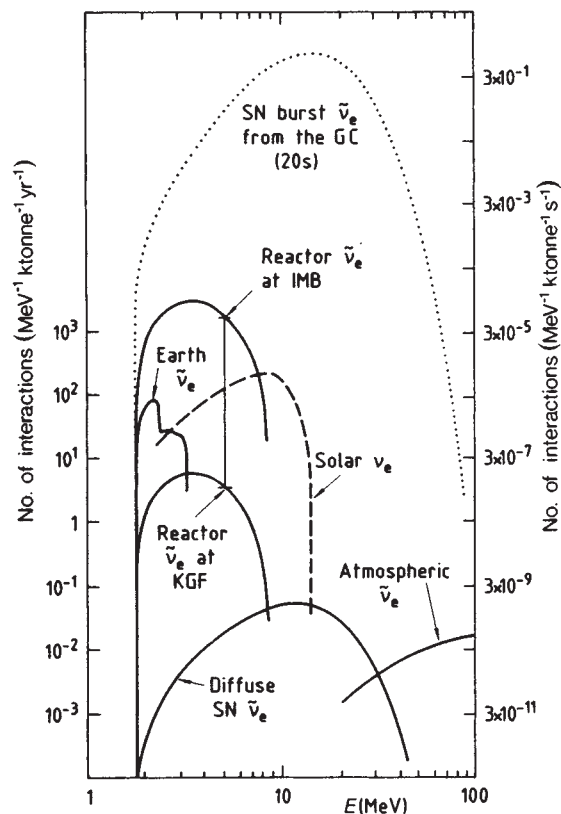


Fig. 1 Number of $\bar{\nu}_e$ interactions expected (through charged current³) $\text{MeV}^{-1} \text{ yr}^{-1}$ (or $\text{MeV}^{-1} \text{ s}^{-1}$), from various sources, in a 1-ktonne scintillator detector (9.4×10^{31} free protons) with 100% efficiency. The shape of the energy spectrum of the reactor $\bar{\nu}_e$ is from ref. 7; the vertical bar indicates the variation of the flux of these $\bar{\nu}_e$ according to the location of the detector (KGF = Kolar Gold Field); the reactors have been assumed to be 100% efficient. The shape of the spectrum of the supernova (SN) $\bar{\nu}_e$ burst is from ref. 8 and that of the diffuse supernova $\bar{\nu}_e$ from ref. 9; the SN burst has been assumed to occur in the galactic centre (GC). Atmospheric $\bar{\nu}_e$ have been taken at Kamioka during minimum solar activity¹⁰. For completeness, the ^8B solar ν_e interactions (through coherent scattering¹¹) in a 1-ktonne scintillator detector (3.5×10^{32} electrons) are shown; the flux has been normalized to the observational data¹².

antineutrino background will be significantly reduced and a threshold of ~ 4 or 5 MeV is predicted; then, the detection of the ^8B solar ν_e by coherent scattering reactions ($\bar{\nu}_e + e^- \rightarrow \bar{\nu}_e + e^-$) will be expected. In these conditions, the reactor $\bar{\nu}_e$ could soon be detectable from several of the present underground laboratories.

Part of my interest in the low-energy antineutrinos originated through an interdisciplinary discussion on the links between astrophysics, particle physics and nuclear physics organized by J. Audouze. I thank all the members of the group and particularly Y. Declais and S. Jullian for stimulating discussions. I also thank J. C. Le Ralle for providing information on the location of nuclear power stations.

Received 24 November 1984; accepted 29 May 1985.

1. Badino, G. *et al. Proc. Neutrino 84 Conf.*, Dortmund, 556–567 (World Scientific Publishing, 1984).
2. Burrows, A. *Astrophys. J.* **283**, 848–852 (1984).
3. Boehm, F. & Vogel, P. *Ann. Rev. nucl. Part. Sci.* **34**, 125–153 (1984).
4. Cavaignac, J. F. *et al. Phys. Lett.* **148B**, 387–394 (1984).
5. Krauss, L. M., Glashow, S. L. & Schramm, D. N. *Nature* **310**, 191–193 (1984).
6. *Nuclear Power Reactors in the World*, (Ref. data Ser. No. 2, IAEA, Vienna, Austria, 1984).
7. von Feilitzsch, F., Hahn, A. A. & Schreckenbach, K. *Phys. Lett.* **118B**, 162–166 (1982).
8. Nadezhin, D. K. & Otroshenko, I. V. *Soviet. Astr.* **24**, 47 (1980).
9. Bisenovaty-Kogan, G. S. & Seidov, Z. F. *Ann. N.Y. Acad. Sci.* **319**–327 (1984).
10. Gaissner, T. K. & Stanev, T. *Proc. Neutrino 84 Conf.*, Dortmund, 370–374 (World Scientific Publishing, 1984).
11. Kayser, B. *et al. Phys. Rev.* **20D**, 87–98 (1979).
12. Davis, R., Cleveland, B. T. & Rowley, J. K. *AIP Conf. Proc.* **96**, 2–15 (1983).

Periodic behaviour of solar flare activity

K. Ichimoto, J. Kubota, M. Suzuki*, I. Tohmura
& H. Kurokawa

Kwasan and Hida Observatories, University of Kyoto, Yamashina,
Kyoto 607, Japan

The periodic nature of solar activity has been studied using parameters such as the sunspot Wolf numbers, calcium plage areas and flare indices. The magnitude of the solar activity based on these parameters reveals periodicities other than the most pronounced 11-yr one. Any absolute detection of periodicity in active phenomena would have fundamental significance for our understanding of solar activity. Here we investigate the temporal variation of the flare activity of the Sun using the data of 8,821 H_{α} flares which occurred during the period January 1965 to February 1984, and show new evidence for 155-day and 17-month periodicities of the flare activity. The 155-day periodicity is examined by taking into account the location of the flare on the Sun. It is suggested that the 155-day period may be related to the timescale for the storage and/or the escape of the magnetic field.

The sunspot Wolf number has been commonly used to investigate the temporal variation of solar activity, but the frequency of flare occurrence is also an excellent indicator. We have studied recently the time change of the spatial distribution of flare occurrences on the Sun during solar cycle 21, and compared it with that of magnetic fields on the solar surface using chronologically arranged magnetic synoptic charts (I.T. *et al.*, in preparation). Many flares occurred just after the appearance of a new activity complex, or within a few rotations of the start of sudden growth of the complex. Our results agree fairly well with those of several other authors^{1,2}. Flare occurrence, therefore, seems to be related to the emergence of the magnetic field on the solar surface. In fact, there is good evidence from optical observations that flares are closely related to the emerging magnetic flux^{1,3,4}. These facts strongly suggest that flare frequency would provide a sensitive indicator of the emerging process of magnetic fields.

Knoška and Křivský⁵ examined the time variation of the occurrence frequency of flares for solar cycle 20. They noted that, in the northern hemisphere, the flare activity was dominant and occurred before that in the southern hemisphere. Recently, Rieger *et al.*⁶ found a significant periodicity of 154 days for the occurrence of γ -ray events observed by the Solar Maximum Mission experiment during February 1980 to September 1983. Any regularities in the occurrence of the flares would provide important clues for understanding the basic mechanism of solar activity.

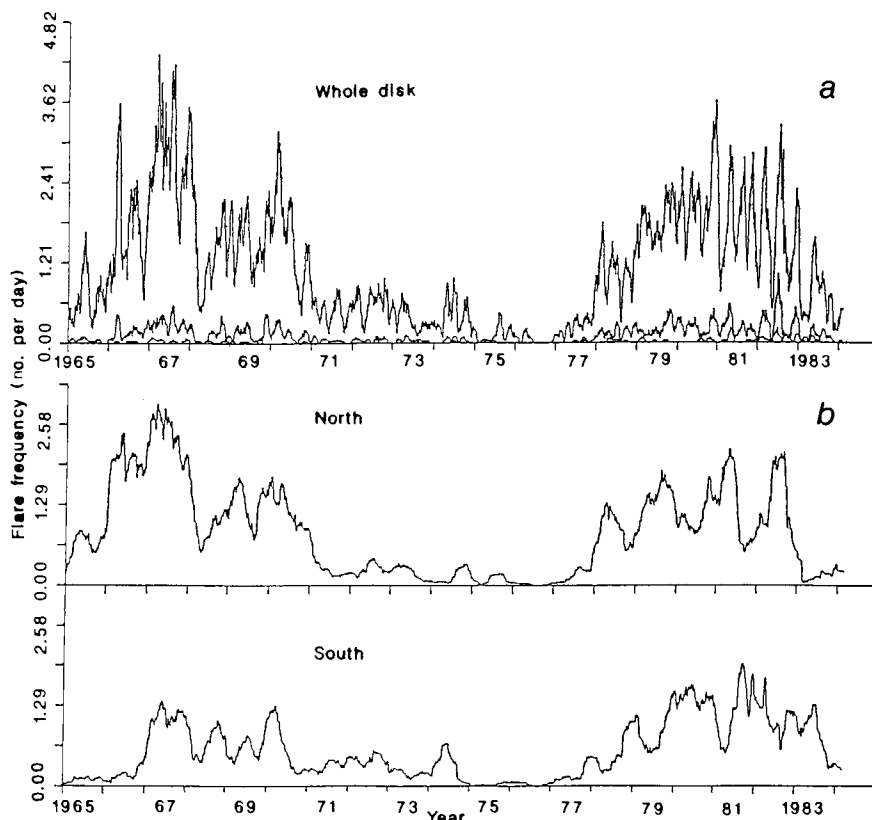
We performed detailed analysis on flare activity using H_{α} flare data of Importance 1, 2 and 3 which occurred during solar cycles 20 and 21. The data used here were taken from ref. 7 for the period from January 1965 to December 1980 and from ref. 8 for January 1981 to February 1984. During solar cycles 20 (1965–75) and 21 (1976–84), 8,821 such flares were recorded.

Figure 1a demonstrates the time variation of a 41-day moving average—that is, $\bar{f}(I) = \sum_{j=-20}^{20} f(I+j)/41$ —for the daily number of flares observed over the whole disk of the Sun for each Importance throughout both solar cycles. This shows a remarkably periodic behaviour of flare occurrence with a period of about 5 months for the declining phase of solar cycle 21. Figure 1b gives the sum of the weighted daily number of flares for the northern and southern hemispheres, where the daily numbers of Importance 1, 2 and 3 are weighted by factors 1, 2 and 3, respectively, and the curves are smoothed by a 141-day moving average. A periodic variation of the flare occurrence with a period of ~ 17 months is found for both hemispheres in solar cycle 21. The anti-correlation between the activities of two hemispheres for solar cycle 21 can also be seen. Figure 1b shows that there seems to be no such anti-correlation between the activities of the two hemispheres during cycle 20. According to Knoška and Křivský⁵, however, the flare activities during solar cycle 20 are enhanced alternately in the northern and southern hemispheres when the occurrence of subflares is taken into account.

Maximum entropy power spectra of the time variation of the weighted flare numbers for solar cycles 20 and 21 allow the periodicities of flare activity to be examined. Figure 2a, b shows the power spectra for the whole disk. A distinct peak is at the period near 155 days in the spectra of both cycles. Another

* Permanent address: Yokkaichi High School, Mie, Japan.

Fig. 1 a, Daily number of H_{α} flares on the whole disk of the Sun for the period January 1965–February 1984 (a 41-day moving average). The top, middle and bottom curves indicate the daily numbers of flares of Importance 1, 2 and 3 respectively. b, The sum of the weighted daily number ranks of 1, 2 and 3 for the northern (upper) and the southern (lower) hemispheres. The daily numbers of flares of Importance 1, 2 and 3 are weighted by factors 1, 2 and 3 respectively.



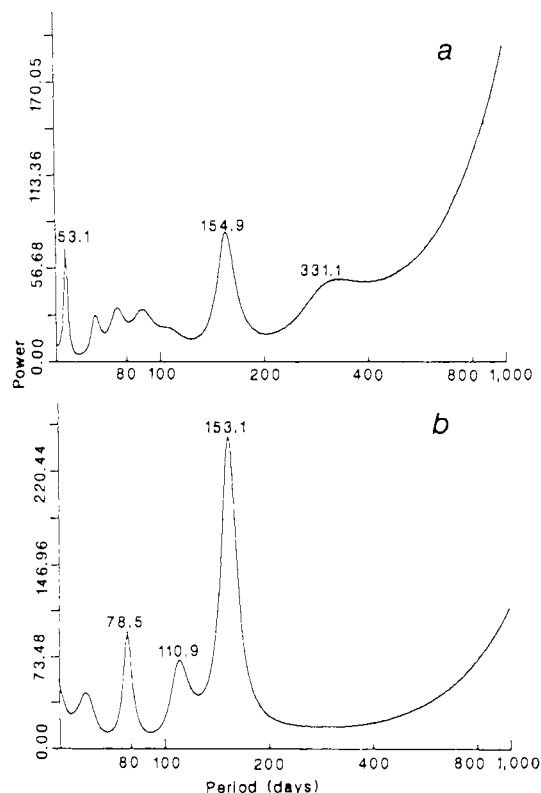


Fig. 2 *a*, The power spectrum of the daily number of flares for solar cycle 20. $N_{\max}$ indicates the length (in days) of the time span for which the power spectrum was calculated: *a*, $N_{\max} = 4,077$; *b*, $N_{\max} = 2,863$. *b*, The same as *a*, but for solar cycle 21.

distinct peak at 510–540 days can also be seen when the power spectra are calculated for northern and southern hemispheres separately for both cycles. These 155- and 520-day periodicities, which were previously predicted for solar cycle 21 from Fig. 1*a*, *b*, are considered to be inherent characteristics of flare activity. We now investigate the nature of the 155-day periodicity.

Rieger *et al.*⁶ found a strong 154-day periodicity in the γ -ray

and large soft X-ray events during the period February 1980–September 1983. During this time, the occurrence of H_{α} flares also shows remarkable periodicity (Fig. 1*a*). The data of flares in this time interval therefore, provide us with a good sample for studying the 155-day periodicity.

To clarify the cause of the 155-day periodicity, the spatial distribution of flares on the solar surface be investigated. Figure 3*a* plots the Carrington longitude of flares which occurred in the 5° N–15° N latitude zone against the rotation number (year). As the period 1980–83 is in the declining phase of solar cycle 21, most flares occurred in this lower-latitude zone. More than half of the flares are concentrated in clusters.

The time-longitude distribution of flares can be compared with that of magnetic fields on the solar surface. The Solar Magnetic Field Synoptic Chart⁸ can be sliced into 5° strips of latitude, and the strips of each latitude zone arranged in chronological order. Figure 3*b* shows the time sequence of the longitudinal distribution of the magnetic flux for the rotations 1,624–1,727 in the latitude zone 10° N–15° N. Many features of magnetized regions, lasting for several rotations, can be easily recognized. Some very long-lasting features exist which drift eastward with time due to the differential rotation. In accordance with Bumba and Howard⁹, we shall call these features streams. Comparison between Fig. 3*a* and *b* reveals that several clusters of flares occur repeatedly at different epochs in a stream.

We have investigated the temporal behaviour of the flare activity with regard to each stream. Two of these in the northern hemisphere are indicated by S1 and S2 in Fig. 3*a*. Figure 4*a* and *b* shows the time change of the daily number of flares that occurred in S1 and S2, respectively; their power spectra are given in Fig. 4*c* and *d*. An obvious peak at 150–160 days can be recognized. Similarly, the power spectra for the streams in the southern hemisphere also show a peak at ~155 days. These results indicate that each stream has a tendency to produce many flares repeatedly with a period of ~155 days, and suggest that this is probably the origin of the 155-day periodicity found in the present study and that of Rieger *et al.*⁶

Note that the time interval 1980–83 also shows the significant 155-day periodicity for the occurrence of flares on the whole disk. If the periodic enhancements of flare activity in each stream are out of phase, the flare occurrence on the whole disk would not have revealed the prominent periodicity seen in Fig. 1*a*. It

Fig. 3 *a*, The time-longitude distribution of flares in the latitude zone 5° N–15° N for the period January 1976 to February 1984. The abscissa is the Carrington longitude, and the ordinate the time (Carrington rotation number). Strips S1 and S2 indicate the longitudinal ranges of long-lasting magnetic features (*b*) in which the temporal behaviour of the flare activity is examined. *b*, The time-longitude distribution of the strongly magnetized region in the latitude zone 10° N–15° N. This chart was reproduced from the Kitt Peak National Observatory Magnetic Synoptic Chart⁸. Many features of the magnetized regions which last for several rotations or more can be seen.

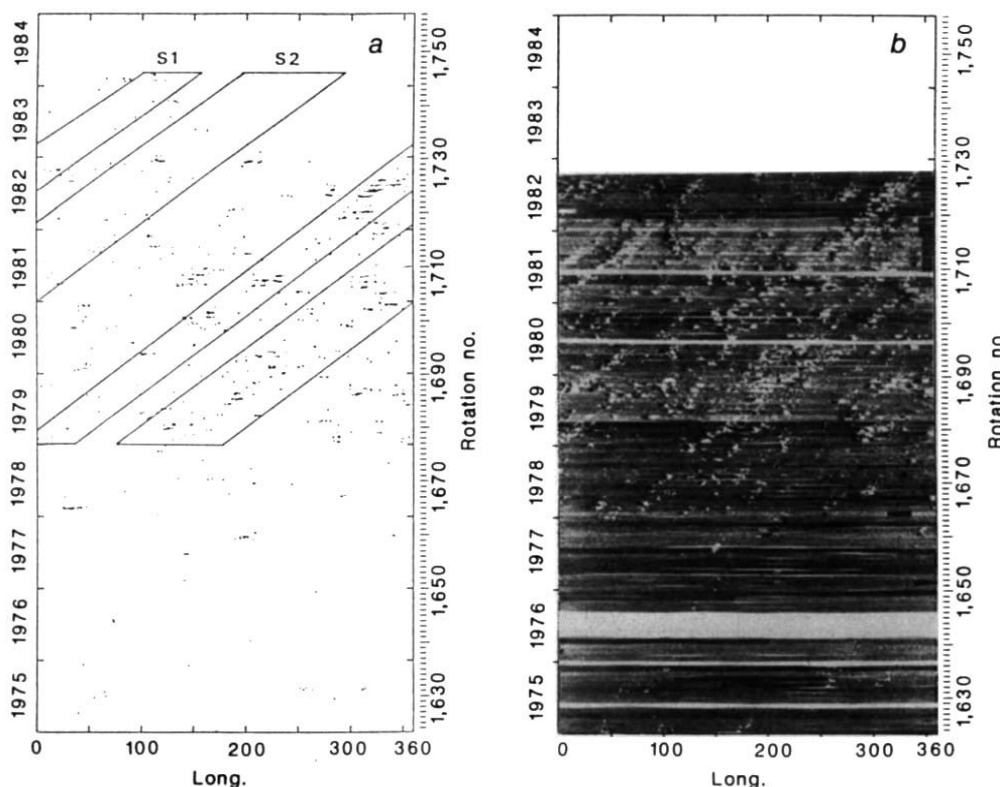
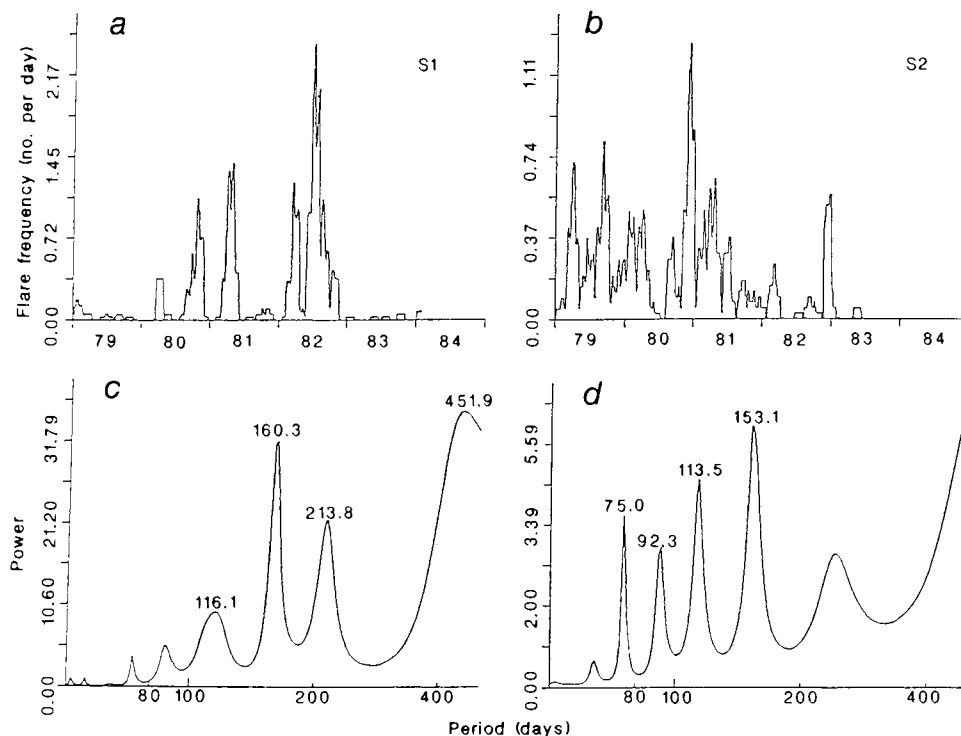


Fig. 4 *a*, The time change of the daily number of flares which occurred in the longitudinal range indicated by strip S1 in Fig. 3*a*. *b*, The same as *a*, but for strip S2. *c*, The power spectrum of the daily number of flares which occurred in the longitudinal range S1, $N_{\text{max}} = 1,858$. *d*, As *c*, but for strip S2, $N_{\text{max}} = 1,858$.



seems, therefore, that the clusters of flares tend to occur simultaneously at different longitudes, at least in this time interval. It is not clear, however, whether or not this is coincidental.

Thus, 155- and 520-day periodicities have been established for H_{α} flares in solar cycles 20 and 21. These periodicities may be regarded as inherent properties of solar flare activity. What of the other indicators of solar activity? The temporal variation of the sunspot Wolf number has been most widely studied¹⁰⁻¹³. One of the most extensive analyses was carried out by Wolff¹³ using monthly mean sunspot number for the time interval 1749-1979. After removing the 11-yr component from the time sequence of the sunspot number, he calculated the Fourier spectrum, and found a significant peak at 155.3 days. This is very close to the period found in the flare occurrence. However, the periodicity of 155 days has not generally been noticed in other work on sunspot number. Perhaps flare occurrence is more sensitively related to the emergence of the magnetic field on the solar surface than to the sunspot number. On the other hand, Yacob and Bhargava¹⁰ found a 16.3-month periodicity both in sunspot number and in the geomagnetic data over 61 yr. This is quite close to our finding of a 17-month periodicity which we shall discuss elsewhere.

The 155-day periodicity of flare occurrence in solar cycle 21 originates in each stream; strongly magnetized regions, which last for a long time and drift in longitude due to the differential rotation, have a tendency to produce flares with a period of ~ 155 days. As mentioned above, flare occurrence can be regarded as the manifestation of the emergence of the magnetic field on the solar surface. It seems, therefore, that the period of 155 days is related to the timescale for the storage and/or escape of the magnetic field in the solar convection zone.

Received 15 March; accepted 5 June 1985.

1. Švestka, Z. in *Solar Flares*, 31-35 (Reidel, Dordrecht, 1976).
2. Gaizauskas, V. & McIntosh, P. S. *Bull. Am. astr. Soc.* **15**, 697 (1983).
3. Kattenberg, A. et al. *Sol. Phys.* **88**, 315-327 (1983).
4. Zirín, H. *Astrophys. J.* **274**, 900-909 (1983).
5. Knoška, Š. & Křivský, L. *Bull. astr. Inst. Czech.* **29**, 352-354 (1978).
6. Rieger, E. et al. *Nature* **312**, 623-625 (1984).
7. *IAU Quarterly Bull. Sol. Activities* Vol. 11-22, Pt III, Eruption Chromospheric (Eidgen, Sternwarte in Zürich, 1965-76, Tokyo Astronomical Observatory, 1977-80).
8. *Solar-Geophysical Data*, No. 438-475, Pt I (US Dept. Commerce, Boulder, Colorado, 1981-84).
9. Bumba, V. & Howard, R. *Sol. Phys.* **7**, 28-38 (1969).
10. Yacob, A. & Bhargava, B. N. *J. atm. terr. Phys.* **30**, 1907-1911 (1968).
11. Knight, J. W., Schatten, K. H. & Sturrock, P. A. *Astrophys. J. Lett.* **227**, L153-L156 (1979).
12. Kapoor, S. G. & Wu, S. M. *J. geophys. Res.* **87**, 9-16 (1982).
13. Wolff, C. L. *Astrophys. J.* **264**, 667-676 (1983).

Daily Earth rotation determinations from IRIS very long baseline interferometry

D. S. Robertson*, W. E. Carter*, J. Campbell† & H. Schuh†

* National Geodetic Survey, Charting and Geodetic Services, National Ocean Service, NOAA, Rockville, Maryland 20852, USA

† Geodetic Institute, Bonn University, Nussallee 17, D-5300 Bonn 1, FRG

Classical optical determinations of Earth rotation (UT1) require several weeks of observations to achieve an accuracy of ~ 1.5 ms of time, whereas modern techniques have demonstrated an accuracy of 0.7-0.9 ms on timescales of 1-5 days^{1,2}. Here we have probed the limits of the accuracy and time resolution of very long baseline interferometry (VLBI) determinations of UT1 by conducting a series of 1-h observing sessions using only one baseline. Comparison of the results from the 1-h sessions with corresponding determinations from 24-h multi-baseline sessions establishes that the accuracy of the 1-h determinations is close to 0.1 ms of time. These observations were scheduled as part of the project MERIT (monitor earth rotation and intercompare techniques of observation and analysis³) intensive observing campaign.

At the time of the MERIT intensive campaign, the IRIS (international radio interferometric surveying⁴) project was already operating 24-h VLBI observing sessions every 5 days to monitor polar motion and UT1. To determine UT1 on the intervening 4 days, 1-h observing sessions were scheduled each day using a 6,000-km baseline which is nearly parallel to the Earth's equatorial plane. For such a baseline, the observed VLBI time delay is nearly a diurnal sinusoid, which can be written as:

$$\tau = l \sin(\text{UT1} + K) \cos(\delta) + k_0 + k_1 t \quad (1)$$

where l is the length of the baseline, δ is the declination of the source, k_0 is the difference in the clock offsets at the two observatories, k_1 is the difference in the clock rates, and t is

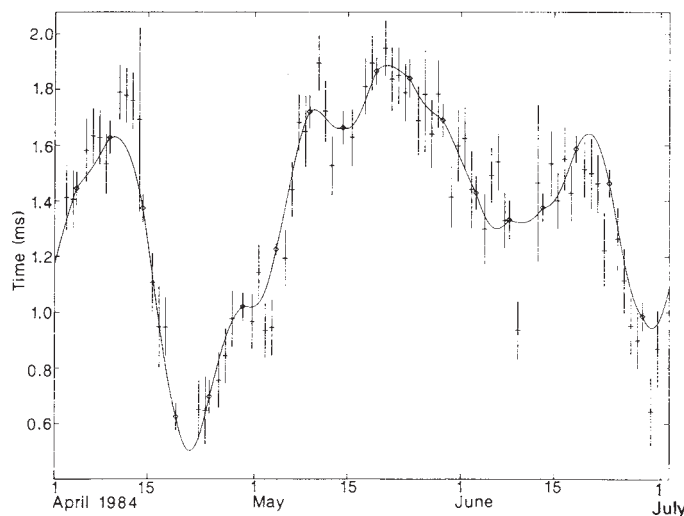


Fig. 1 Determinations of UT1, shown as differences from the BIH UT1R plus the Yoder⁷ short-period tidal corrections. The intensive campaign observations are shown as crosses, and the regular IRIS observations are shown as diamonds. The smooth curve has been fitted to the regular IRIS determinations alone.

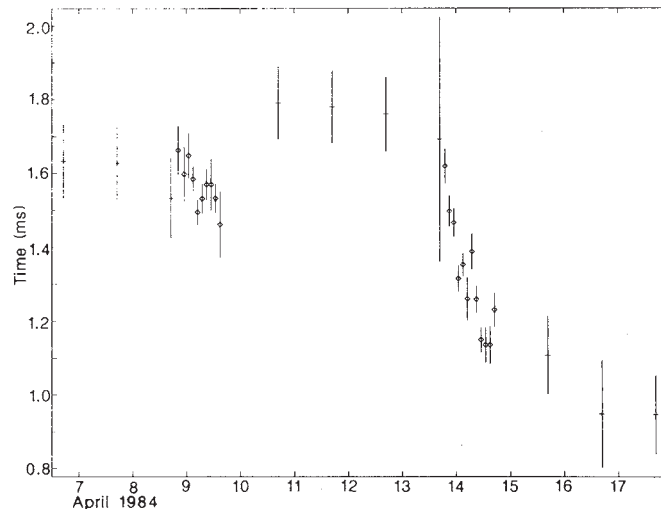


Fig. 2 Results of special analysis of two IRIS regular sessions broken into 2-h subsets (marked with diamonds), along with 10 intensive campaign sessions spanning the period from 7 to 17 April 1984.

time. The phase angle K is the angle between the baseline and the source at some arbitrary origin in time. If we assume that l , K , δ , k_0 and k_1 are known from previous observations, then a single observation of τ for a low-declination source would be sufficient for determining UT1. However, as the clocks are not perfect, and k_0 and k_1 must also be determined from the observations, a set of observations which span a range in either UT1 or δ is needed.

The baseline selected for the IRIS daily observations involved the Westford observatory in Massachusetts (USA) and the Wettzell observatory in Bavaria (FRG). This baseline was selected principally because it is nearly parallel to the Equator (its declination is only 5°) and its length is $\sim 6,000$ km. The use of a shorter baseline would needlessly sacrifice sensitivity to UT1, while the use of a longer baseline would create problems in obtaining sufficient observations at high enough elevation angles to minimize atmospheric effects. The schedules selected included four sources at declinations ranging from 13° to 73° . Each source was observed twice to provide redundancy in case of the loss of any given observation. The observing schedule was repeated at the same sidereal time each day for the 3-month period, except for those days in which the Westford and Wettzell observatories were occupied with the normal IRIS observations (every fifth day). Table 1 shows a sample of the daily observing sequence.

One-hour observing sessions were scheduled on a total of 73 days during April, May and June of 1984; 64 of these sessions produced accurate UT1 values. The regular IRIS observations at 5-day intervals normally involved four stations (Westford, Wettzell, plus the George R. Agassiz station in Texas and the Richmond observatory in Florida) for 24-h observation periods.

Table 1 Sample IRIS intensive campaign observing schedule

yr	month	day	h	min	Source
1984	04	01	17	30	0234+285
1984	04	01	17	39	0212+735
1984	04	01	17	48	0552+398
1984	04	01	17	57	0528+134
1984	04	01	18	11	0234+285
1984	04	01	18	20	0212+735
1984	04	01	18	28	0552+398
1984	04	01	18	37	0528+134

The four quasars that were selected span 60° in declination and were all above 30° elevation at both sites during the observation periods.

Three of the sessions also included the Onsala Space Observatory in Sweden.

The formal errors for the UT1 determinations were based on observation weights adjusted so that the χ^2 per degree of freedom of the post-fit residuals was close to unity. The average formal error was 0.1 ms for the intensive campaign observations, and 0.05 ms for the regular IRIS observations.

To minimize atmospheric effects, all of the observations in the intensive campaign schedule were taken at elevation angles of 30° or higher at both stations. These observations were processed using a fixed atmosphere model based on surface meteorological measurements at the stations. To test the sensitivity of the resulting UT1 determinations to errors in this atmosphere model, test solutions were run in which an atmosphere zenith parameter was estimated at each station. In a typical case, the adjustments to the atmosphere zenith thickness were 0.15 ± 0.24 ns at Westford and 0.11 ± 0.14 ns at Wettzell. Adding these parameters to the solution only changed the estimated UT1 value by 0.05 ms of time, about half the formal error of the UT1 estimate.

Because the observing sessions were only about 1 h in duration, we did not expect to find that clock fluctuations would introduce serious errors. To test this assumption, we estimated a parabolic clock parameter in a sample case. The addition of this parameter to the solution changed the estimated UT1 value by only 0.003 ms of time.

The major systematic error source in these intensive campaign UT1 determinations probably results from errors in the assumed pole positions. Numerical experiments indicate that the effect of a pole position error will be ~ -0.03 ms of time per ms of arc in X and 0.05 ms of time per ms of arc in Y . The pole values used to reduce the intensive campaign observations came from the routine IRIS observing sessions performed at 5-day intervals, as published in IRIS Bulletin A⁵. Intercomparisons of these pole values with corresponding determinations from satellite laser observations places limits on their errors at the level of ± 2 arc ms (ref. 6), so the UT1 errors which result from pole position errors should not be any larger than the formal errors of the UT1 determinations.

The redundancy built into the observing schedules to protect against the loss of a single observation (that is, the repetition of the sequence of four observations) provided a useful check of the internal consistency of these determinations. A sample of 10 observing days was selected in which all eight observations were successful, and these days were reprocessed to determine separate values of UT1 for each of the two sets of observations

of the four sources. The weighted root mean square (r.m.s.) difference between the two UT1 determinations on those days was only ~ 0.1 ms, which is about the size of the formal error of the estimates.

The UT1 determinations from the intensive campaign observations are shown in Fig. 1, along with the regular IRIS five-day observations during the same time period. To determine quantitatively how well the two series of observations agree, a curve was interpolated through the five-day IRIS observations alone, using a $\sin(t)/t$ interpolating function. The weighted r.m.s. difference between the intensive campaign observations and the values obtained from this curve was only 0.1 ms of time, in good agreement with our estimates for the errors in these determinations. This agreement between the two different sets of UT1 determinations places a strong bound on the total errors in both series, in the absence of common sources of systematic error. The likelihood of such common error sources is small because of the differences in the observation schedules and the number of baselines used.

Another way of testing the validity of the 1-h intensive campaign observations is to compare these with the trends exhibited by the regular 24-h IRIS observing sessions. If the data from the 24-h sessions are divided into smaller subsets which are then used to estimate UT1, trends are revealed which should agree with those exhibited by the intensive campaign observations. A typical example comprising two IRIS experiments, each broken into 2-h subsets, is shown in Fig. 2. The 8–9 April values cluster within a total range of only 0.2 ms, and agree well with the relatively gentle variations of the intensive series over a period of several days spanning that session. In contrast, the 13–14 April values display a significant linear trend, indicating a change in UT1 of 0.5 ms during the 24-h period. When compared with the intensive series, the agreement could hardly be better. The 13–14 April IRIS session fortuitously captured a good fraction of a change in UT1 of more than a full millisecond which occurred in a period of only 10 days.

The sizable number of one-day data points obtained during the three months of the MERIT intensive campaign gives us an opportunity to examine the short-period spectrum of the UT1 fluctuations. Variations with periods between five and 35 days caused by tidal deformation of the polar moment of inertia have been derived theoretically, for example, by Yoder *et al.*⁷. The periods and phases of these components of Earth rotation are precisely determined by the motions of the Moon and the Sun, but the amplitudes are uncertain because of imperfections in the model for the tidal deformations of the fluid core and the oceans.

A preliminary spectral analysis revealed sharp peaks at the frequencies expected from the tidal analysis. To analyse the short-period spectrum in detail we began with the total UT1–UTC values (as tabulated in IRIS Bulletin A⁵), and subtracted the contribution expected from meteorological effects calculated from the global wind data determined by the United States

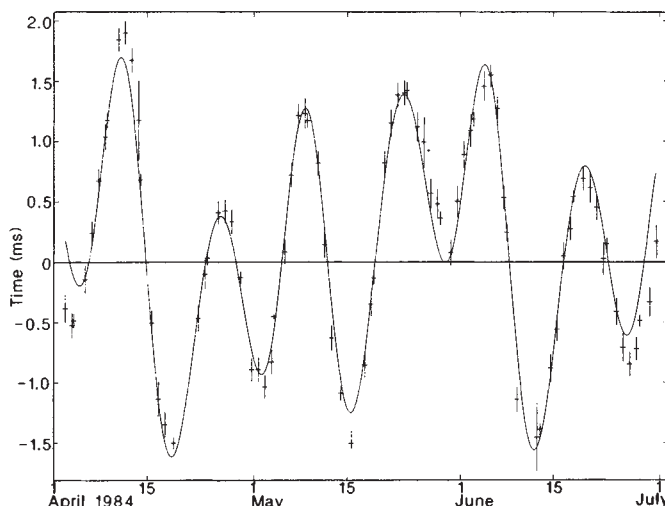


Fig. 3 VLBI estimates of UT1–UTC with the atmospheric contribution and a best-fit parabola removed (see text). The smooth curve represents the five superposed sine waves whose coefficients are listed in Table 2.

National Meteorological Center⁸. We then subtracted a best-fit parabola to remove long-term drifts. The resulting data are shown in Fig. 3. We then fitted four sinusoids to the data shown in Fig. 3. The periods of the sinusoids were fixed at the values of the largest terms < 40 days in Yoder's model (9.13, 13.66, 27.56 and 31.81 days). The residuals to this fit were found to have a smooth drift with a period of ~ 60 days, apparently resulting from the inaccuracy of the weather data. (There are no significant tidal terms between 35 and 90 days.) Therefore, we added one more sinusoid to our fit, and adjusted its period near 60 days. The periods and phases of the sinusoids resulting from this fit are listed in Table 2 along with the expected amplitudes and phases from the theoretical work. The smooth curve in Fig. 3 was calculated from these sinusoids. The difference between the sinusoidal model and the VLBI determinations has an r.m.s. value of only 0.17 ms, and a large fraction of this difference probably results from remaining errors in the atmosphere observations. The 0.43-ms amplitude of the 60-day sinusoid represents an error of only 0.03 ms day^{-1} in the atmosphere data, and previous studies have indicated that the error in the atmosphere data could be of the order of 0.1 ms day^{-1} (ref. 2). The uncertainty in the wind data places a fundamental restriction on our ability to use accurate measurements of UT1 to determine, for example, the scaling parameter k/C used to calculate the tidal corrections.

The IRIS observations have demonstrated that a single VLBI baseline, observing for only 1 h, can determine UT1 with a precision of 0.1 ms of time. Intercomparison with determinations from 24-h VLBI sessions using more baselines has established

Table 2 Results of preliminary frequency analysis applied to the daily UT1 values determined by VLBI

Theoretical			Observed		
Period (days)	Amplitude (ms)	Phase* (deg)	Period (days)	Amplitude (ms)	Phase* (deg)
9.12	-0.04	289.8			
9.13	-0.10	359.4	9.13	-0.10 ± 0.01	360.8 ± 3.9
13.63	-0.32	38.5			
13.66	-0.78	108.3	13.66	-1.05 ± 0.01	81.4 ± 0.7
27.09	0.04	216.8			
27.44	0.05	181.7			
27.56	-0.83	251.2	27.56	-0.75 ± 0.02	247.2 ± 1.5
27.67	0.05	320.8			
31.81	-0.18	190.8	31.81	-0.30 ± 0.02	127.8 ± 3.6
			61.8 $\pm$ 0.8	0.43 ± 0.01	137.3 ± 3.4

* Phase is referred to 4 April 1984, 12 h UT.

that the accuracy of the determinations is probably no worse than 0.1–0.2 ms, given that the two types of observing sessions do not contain common sources of systematic error. The success of these very brief VLBI observing sessions opens the possibility of continued inexpensive UT1 determinations and points the way to further exploring the spectrum of UT1 fluctuations on very short timescales. The daily observations will resume in April 1985 and continue for at least 1 yr. This continuing data set should enable the spectrum of UT1 fluctuations to be reliably determined down to periods of ~ 2 days.

Space restrictions prevent us from acknowledging the contributions of each of the many organizations and individuals that have contributed to the IRIS project. However we must thank the Special Research Group on Satellite Geodesy (SFB 78) and the staff of the Wettzell Observatory, as well as the Northeast Radio Observatory Corporation (NEROC) and the staff of the Westford Observatory for their exceptional efforts in support of the MERIT Intensive Campaign.

Received 22 March; accepted 10 June 1985.

1. Robertson, D. S. *et al.* *Nature* **302**, 509–511 (1983).
2. Carter, W. E. *et al.* *Science* **224**, 957–961 (1984).
3. Wilkins, G. A. (ed.) *Project MERIT: A Review of the Techniques to be Used During Project MERIT to Monitor the Rotation of the Earth* (Royal Greenwich Observatory, Herstmonceux, 1980).
4. Carter, W. E. & Robertson, D. S. *Proc. int. Symp. on Space Techniques for Geodynamics*, Sopron, Hungary (1984).
5. *IRIS Bull. A* (National Geodetic Survey, NOAA N/CG114, Rockville, Maryland, 1984).
6. Robertson, D. S., Carter, W. E., Tapley, B. D., Schutz, B. E. & Eanes, R. J. *Polar Motion Measurements: Sub-Decimeter Accuracy Verified by Intercomparison*, Preprint (NOAA, Rockville, Maryland, 1985).
7. Yoder, C. F., Williams, J. G. & Parke, M. E. *J. geophys. Res.* **86**, 881–891 (1981).
8. Rosen, R. D. & Salstein, D. A. *J. geophys. Res.* **88**, 5451–5470 (1983).

Dissolved aluminium in the central North Pacific

Kristin J. Orians & Kenneth W. Bruland

Center for Marine Studies and Board of Study in Chemistry,
University of California, Santa Cruz, California 95064, USA

Aluminium is the most abundant metallic element in the Earth's crust (8.23% by weight)¹, yet little is known about its oceanic distribution. Published data sets concerning aluminium in seawater² are primarily for the North Atlantic Ocean^{3–7}. We report here that dissolved aluminium concentrations in the central North Pacific are 8–40 times lower than those at corresponding depths in the central North Atlantic, but the vertical distribution features are similar. The vertical distribution and inter-ocean fractionation of aluminium can be explained by geographical variations in atmospheric aluminium sources, intense particle scavenging throughout the water column, and some regeneration in bottom waters. Aluminium's short oceanic residence time (estimated here as 100–200 years) leads to its marked inter-ocean fractionation, which is the reverse of that for nutrient elements such as silicon.

Aluminium(III) has a strong tendency to hydrolyse in water. At the pH of seawater (7.5–8.3), $\text{Al}(\text{OH})_3^0$ and $\text{Al}(\text{OH})_4^-$ are predicted to be the dominant forms of dissolved aluminium². The hydrolysis products of trivalent and tetravalent metals, such as Al, Fe, and Th are particle-reactive species with relatively short oceanic residence times. Available data suggest that the concentration of dissolved aluminium in seawater is lower than predicted by thermodynamic equilibria with various aluminosilicates^{8,9}. This is not unexpected, however, because the concentrations of many trace elements are apparently controlled by scavenging processes rather than solubility with a solid phase¹⁰. Possible removal mechanisms for aluminium include passive adsorption onto particle surfaces or active biological uptake, either as a micronutrient or a component of siliceous skeletal material. The particle types responsible for the passive adsorption/scavenging of dissolved aluminium may include inorganic (such as biogenic opal, clay minerals, or iron and manganese oxyhydroxides), organic, and inorganic substrates with an

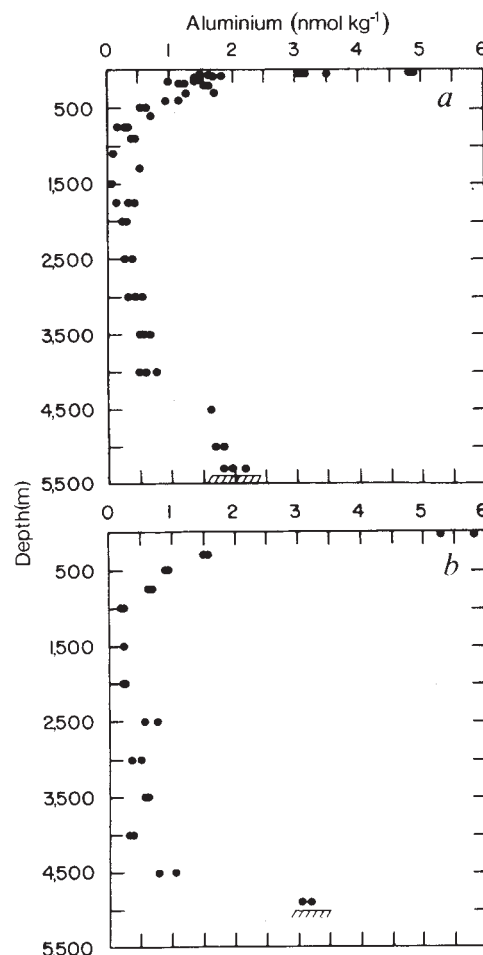


Fig. 1 Variation of dissolved aluminium with depth at: *a*, 28° N, 155° W (VERTEX IV, July 1983); *b*, 14° N, 160° W (Marine Chemistry, station 17, October 1980). Concentrations were determined by liquid-liquid extraction with 8-hydroxyquinoline into chloroform, followed by HGA atomic absorption analysis. Samples were collected and filtered as described previously^{31,32}, then stored frozen at natural pH for up to 4 yr. Analytical precision (per cent difference for duplicate sample analysis) was 5% for samples with dissolved aluminium concentrations $> 1.0 \text{ nmol kg}^{-1}$ and averaged 11% overall. Blanks, determined by extracting quartz distilled water, averaged $0.15 \text{ nmol kg}^{-1}$, and the detection limit (2 s.d. for blanks) was $0.06 \text{ nmol kg}^{-1}$. Details of the technique will be presented elsewhere.

organic coating¹¹. The North Atlantic open ocean data are consistent with passive adsorption as the dominant removal mechanism^{3,4}, while evidence from the Mediterranean, a confined marine basin, and laboratory studies^{12–14} have suggested an active biological uptake removal mechanism.

In this study, the distribution of dissolved aluminium in the North Pacific was investigated to elucidate the oceanic chemistry of this element. Profiles of dissolved aluminium from the North Pacific, together with data from the North Atlantic, would provide a comparison of ocean basins with markedly different characteristics. The North Atlantic is influenced to a greater extent by the atmospheric and fluvial input of crustal material than is the North Pacific, due to its smaller size and enhanced sources. In addition, the North Atlantic deep waters are recently formed, and thus less influenced by the nutrient cycles which tend to enrich greatly the older North Pacific deep waters with elements associated with biological cycles.

Vertical distributions of dissolved aluminium were determined at two locations in the central North Pacific; 28° N, 155° W (VERTEX IV, July 1983), and 14° N, 160° W (Marine Chemistry, station 17, October 1980). Both areas showed a surface maximum of 4–6 nmol kg^{-1} , a broad minimum of $< 0.5 \text{ nmol kg}^{-1}$ between 750 and 3,000 m, and an increase with depth below 3,000 m to 2–3 nmol kg^{-1} near the bottom (Fig. 1). These concentrations

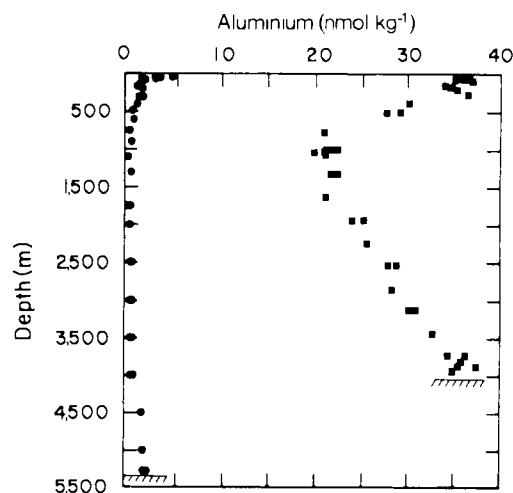


Fig. 2 Variation of dissolved aluminium with depth at 28° N, 155° W in the North Pacific (●) and at 41° N, 64° W in the North Atlantic (■). The Pacific data are from VERTEX IV and the Atlantic data are from K69/10, by Hydes³. Copyright 1979 by the AAAS. A deep western North Atlantic seawater reference sample (NASS-1) was analysed by our technique and found to have a dissolved aluminium concentration comparable with that Hydes³ found at the same depth at his western North Atlantic station.

are significantly lower than those published for the central North Atlantic³. The VERTEX IV profile from the North Pacific is plotted together with Hydes' North Atlantic data in Fig. 2 to emphasize the contrast in dissolved aluminium between the two ocean basins. At depths of 1,000–2,000 m, corresponding to the dissolved aluminium minimum, the North Pacific values are roughly a factor of 40 lower than those in the North Atlantic.

The surface concentration of dissolved aluminium should be controlled by several factors: (1) the aeolian flux of particulate aluminium together with the fraction of the particulate aluminium which dissolves in the surface water; (2) the atmospheric flux of dissolved aluminium in rain during wet deposition; (3) the amount of dissolved aluminium from fluvial sources which reaches the open ocean and (4) the rate at which the dissolved aluminium is scavenged from the surface waters. The atmospheric flux of particulate aluminium has been estimated at several mid- and low-latitude locations in the North Atlantic and North Pacific oceans. The estimates are summarized in Table 1. Although these estimates vary widely, the aluminium flux to the Atlantic ocean is roughly $10,000 \text{ ng cm}^{-2} \text{ yr}^{-1}$ compared with estimates for the Pacific of $\sim 1,500 \text{ ng cm}^{-2} \text{ yr}^{-1}$. The Sahara Desert has a major influence on the dust flux in the mid- and low-latitude regions of the North Atlantic^{15,16}. This dust input has been shown to be a factor in the variability of surface concentrations of dissolved aluminium within the North

Atlantic⁴. The roughly seven-fold difference in the estimates of atmospheric flux between the North Pacific and the North Atlantic is similar to the eight-fold difference observed in dissolved aluminium concentrations between the surface waters of the two oceans (Fig. 2).

Little information is available concerning the fraction of the total atmospheric aluminium flux which is in the dissolved form. Measurements of dissolved and particulate aluminium in rain water in Florida (Prospero *et al.*, personal communication) have shown that generally <5% of the total aluminium is dissolved (median = 2.5%, mean = 8.9%). The fraction of particulate aluminium which dissolves in surface seawater is also poorly known but reported¹⁷ to be <1%. If, as a first approximation, these two factors are assumed to be constant from ocean to ocean, then the observed correlation between dissolved aluminium concentrations in the surface waters of the central gyres and the total atmospheric aluminium flux is reasonable.

The vertical distribution of dissolved aluminium in these open ocean sites supports a passive adsorption removal mechanism. Dissolved aluminium is not found to correlate directly with any nutrient-type element at these open ocean sites, but rather shows an inverse relationship to nutrients in the surface waters and through the thermocline, implying scavenging rather than remineralization as the dominant process controlling the vertical distribution of dissolved aluminium. Nutrient-type elements show an enrichment in the older North Pacific deep waters whereas the reverse has been found for aluminium. Thus, there is a trend of inverse deep-water concentration patterns for silicon and aluminium between the North Atlantic and the North Pacific ocean basins. Deep waters of the Mediterranean sea are an extreme example of this fractionation, with the highest aluminium concentration and the lowest silicon concentration¹².

The profiles of dissolved aluminium (Fig. 1) imply a bottom water source of aluminium. The rapid decrease above the bottom with an intermediate depth minimum, however, also implies substantial *in situ* scavenging of dissolved aluminium from the deep waters. Application of a simple vertical advection-diffusion scavenging model¹⁸ to the dissolved aluminium data at the VERTEX-IV site (2,000–5,000 m) yields a residence time ($T_{1/2}$) with respect to scavenging of 150–200 yr (zero order and first-order treatments give comparable results). This is consistent with the observed differences in deep-water dissolved aluminium concentrations with respect to horizontal advection, considering the transit time of deep water from the North Atlantic to the North Pacific¹⁹. During this transit time, the dissolved aluminium input from the North Atlantic would decrease by *in situ* scavenging from the original concentration of 30–40 nmol kg^{-1} in the deep North Atlantic to the low values observed in the deep North Pacific.

The concentration of dissolved aluminium in the deep water could also reflect the surface concentrations at that site by the redissolution of previously adsorbed aluminium. For example, consider the case where biogenic opal is a carrier phase for scavenging dissolved aluminium from the surface waters^{20,21}. Assuming that similar amounts of opal are produced in the central North Pacific and North Atlantic, then the difference in surface aluminium concentrations could lead to more dissolved aluminium being scavenged and higher aluminium to opal ratios in sinking particles in the North Atlantic than in the North Pacific. During dissolution of this opal carrier phase, more dissolved aluminium would be released into the deep North Atlantic than the deep North Pacific, thus transferring the surface signal to the deep waters.

Rough estimates of mean oceanic residence times can be made by comparing the amount of dissolved aluminium in the world's oceans with the input rates of the two main sources, aeolian and fluvial. Our data suggest that the mean oceanic dissolved aluminium concentration is probably closer to 2 nmol kg^{-1} than the previous estimates of 20–70 nmol kg^{-1} (refs 22, 23). If $\sim 2.5\%$ of a total average atmospheric aluminium particulate flux of $4,000 \text{ ng cm}^{-2} \text{ yr}^{-1}$ (Table 1) dissolves, the residence time ($T_{1/2}$) with respect to aeolian input alone is 150 yr. The residence time

Table 1 Atmospheric flux of total aluminium over the North Atlantic and North Pacific Oceans

Location	Aluminium flux ($\text{ng cm}^{-2} \text{ yr}^{-1}$)	Ref.
Pacific		
0–50° N (average)	1,100–2,220	30
40–50° N	1,760–2,640	30
25–40° N	1,520–3,600	30
15–25° N	1,120–1,840	30
5–15° N	520–1,200	30
0–5° N	104–536	30
Atlantic		
0–50° N (Average from Sahara)	22,800–76,200	28
Average North Atlantic	5,000	29
Eastern Atlantic:		
30–40° N	8,500–15,000	16
10–30° N	51,000–80,000	16
5–10° N	8,500–15,000	16

for dissolved aluminium with respect to scavenging calculated from the deep North Pacific profiles is similar. With respect to the best estimate of riverine input²⁴ the residence time is only 40 yr. However, this approach is not entirely valid due to the fact that dissolved aluminium delivered by rivers may be largely removed in the estuaries²⁵⁻²⁷ and coastal zone, and fluvial input may not be a large net source of dissolved aluminium to the central oceans.

We conclude that the concentration of dissolved aluminium is 8-40 times lower in the central North Pacific than in the central North Atlantic. No other element has been reported to show an enrichment in the Atlantic over the Pacific of this great a magnitude. This interbasin difference seems to be due to the short residence time of dissolved aluminium coupled with the differences in atmospheric input, though more work is needed to understand fully the processes involved. Estimates of the mean oceanic concentration of dissolved aluminium must be lowered from 20-70 to ~2 nmol kg⁻¹. Estimated residence times are all short (<200 yr) supporting the conclusion that dissolved aluminium, probably in the form of Al(OH)₃⁰ and Al(OH)₄⁻, is an extremely particle-reactive element in the ocean. Although there is *in situ* scavenging throughout the water column, regeneration of dissolved aluminium at or near the sediment interface results in increased concentrations in the bottom waters.

We thank Rob Franks and Margaret Delaney for their critical review of this paper, and Geoff Smith for help with sample collection. This research was supported by NSF grant OCE-8216672.

Received 29 January; accepted 5 June 1985.

1. Taylor, S. R. *Geochim. cosmochim. Acta* **28**, 1273-1285 (1964).
2. Bruland, K. W. in *Chemical Oceanography*, Vol. 8 (eds Riley, J. P. & Chester, R.) 157-220 (Academic, London, 1983).
3. Hydes, D. J. *Science* **205**, 1260-1262 (1979).
4. Hydes, D. J. *Geochim. cosmochim. Acta* **47**, 967-973 (1983).
5. Moore, R. M. *Geochim. cosmochim. Acta* **45**, 2475-2482 (1981).
6. Olafsson, J. in *Trace Metals in Seawater* (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 475-485 (Plenum, New York, 1983).
7. Measures, C. I., Grant, B., Khadem, M., Lee, D. S. & Edmond, J. M. *Earth planet. Sci. Lett.* **71**, 1-12 (1984).
8. Stoffyn, M. *Science* **203**, 651-653 (1979).
9. Hydes, D. J. *Nature* **268**, 136-137 (1977).
10. Goldberg, E. D., Broecker, W. S., Gross, M. G. & Turekian, K. K. in *Radioactivity in the Marine Environment*, 137-146 (National Academy of Sciences, Washington DC, 1971).
11. Neihof, R. A. & Loeb, G. I. *J. mar. Res.* **32**, 5-12 (1974).
12. MacKenzie, F. T., Stoffyn, M. & Wollast, R. *Science* **199**, 680-682 (1978).
13. Stoffyn, M. *Science* **203**, 651-653 (1979).
14. Stoffyn, M. & MacKenzie, F. T. *Mar. Chem.* **11**, 105-127 (1982).
15. Schutz, L., Jaenicke, R. & Pietrek, H. *Geol. Soc. Am. Spec. Pap.* **186**, 87-100 (1981).
16. Chester, R. *Mar. Chem.* **11**, 1-16 (1982).
17. Hodge, V., Johnson, S. R. & Goldberg, E. D. *Geochem. J.* **12**, 7-20 (1978).
18. Craig, H. *Earth planet. Sci. Lett.* **23**, 149-159 (1974).
19. Broecker, W. S. & Peng, T. *Tracers in the Sea*, 236-243 (Eldigio, New York, 1982).
20. Lewin, J. C. *Geochim. cosmochim. Acta* **21**, 182-198 (1961).
21. Stoffyn-Elgi, P. *Geochim. cosmochim. Acta* **46**, 1345-1352 (1982).
22. Whitfield, M. *Interdis. Sci. Rev.* **6**, 12-35 (1981).
23. Quinby-Hunt, M. S. & Turekian, K. K. *EOS* **64**, 130-131 (1983).
24. Martin, J. M. & Whitfield, M. in *Trace Metals in Seawater*, (eds Wong, C. S., Boyle, E., Bruland, K. W., Burton, J. D. & Goldberg, E. D.) 265-296 (Plenum, New York, 1983).
25. Hydes, D. J. & Liss, P. S. *Estuar. coast. mar. Sci.* **5**, 755-769 (1977).
26. Scholkovitz, E. R. *Earth planet. Sci. Lett.* **41**, 77-86 (1978).
27. Mackin, J. E. & Aller, R. C. *Mar. Chem.* **14**, 213-232 (1984).
28. Prospero, J. M. in *The Sea*, Vol. 7 (ed. Emiliani, C.) 801-874 (Wiley, New York, 1981).
29. Buat-Menard, P. & Chesselet, R. *Earth planet. Sci. Lett.* **42**, 399-411 (1979).
30. Uematsu, M. *et al. J. geophys. Res.* **88**, 5343-5352 (1983).
31. Bruland, K. W., Franks, R. P., Knauer, G. A. & Martin, J. H. *Analyt. chim. Acta* **105**, 233-245 (1979).
32. Bruland, K. W. *Earth planet. Sci. Lett.* **47**, 176-198 (1980).

Age and palaeoclimatic significance of the loess of Lanzhou, north China

Douglas W. Burbank

Department of Geological Sciences, University of Southern California, Los Angeles, California 90089-0741, USA

Li Jijun

Department of Geology and Geography, Lanzhou University, Lanzhou, Gansu, People's Republic of China

Extending in a broad arc that abuts the sandy (non-Gobi) deserts, the loess plateau of northern China¹⁻³ is one of the most massive accumulations of loess in the world. The loess sequence is typically characterized by an alternation of silty or sandy loess with more clay-rich palaeosols. These alternations, in conjunction with their enclosed faunas and distinctive mineralogies, have been interpreted as reflecting Pleistocene glacial/interglacial cycles³⁻⁸. Because it holds implications for the climatic and anthropological history of China^{9,10}, the definition of a reliable chronological framework for loess deposition is of great interest. Recent palaeomagnetic studies^{4,5} have indicated that loess deposition in Shaanxi province commenced ~2.4 Myr ago. To assess the synchrony of loess accumulation across the loess plateau, we have dated a 330-m-thick loess sequence near Lanzhou, Gansu province. The magnetostratigraphical results reported here indicate that the base of this loess succession dates from ~1.3 Myr ago. This young age (in comparison to the Shaanxi sequence) is attributed to uplift along the northern fringe of the Tibetan Plateau¹¹ that precluded early Pleistocene loess preservation in this mountainous region. Palaeosols in the basal loess occur, on average, once every 25 kyr, suggesting that climates conducive to soil-forming events may have been modulated by orbital precession in the early Pleistocene.

Within the extensive loess plateau (Fig. 1a), the best-documented loess successions are in Shaanxi province^{1,9}, where the loess mantle averages 80-140 m in thickness (Fig. 1b). Recent magnetostratigraphical investigations at Luochuan in Shaanxi^{4,5} have indicated that loess deposition commenced ~2.4 Myr ago,

a date remarkably similar to the timing of both the initiation of loess deposition in Europe^{12,13} and the isotopically-determined onset of major Northern Hemisphere glaciation¹⁴. Given this apparently synchronous response on a global scale to Plio-Pleistocene climatic change, it might be assumed that the loess of north China spans a consistent chronological interval throughout most of its areal extent.

To test this assumption, we have created a magnetostratigraphy for the loess sequence at Lanzhou, 500 km to the west of Luochuan (Fig. 1b). At Lanzhou, the highest fluvial terrace of the Huang Ho is overlain by >330 m of loess and appears to be the oldest loess-mantled surface in the area. Wang's¹⁵ palaeomagnetic studies of this sequence used alternating-field (a.f.) demagnetization, placed the Brunhes/Matuyama boundary 105 m above the base, delineated three normal magnetozones in the late Matuyama chron, and indicated increasingly dispersed magnetic orientations (suggestive of incomplete overprint removal) in strata older than 500 kyr. In the present study, the lowermost 100 m of loess were sampled along the completely exposed walls of a steep gully. In the succeeding 230 m, pits up to 5 m deep were excavated to obtain samples of the undisturbed loess from beneath any colluvial material. Palaeomagnetic sampling sites, comprising three or four specimens each, were spaced at ~7-m intervals and, wherever possible, were chosen so as to avoid palaeosols. Pilot studies using stepwise a.f. and thermal demagnetization indicate that: (1) some reversely magnetized specimens exhibit a soft, viscous overprinting of normal polarity; (2) this overprinting can usually be removed at low demagnetization levels (150 Oe or 200 °C); (3) above these levels, most specimens exhibit a stable characteristic remanence; and (4) the primary magnetic carrier appears to be magnetite or titanomagnetite with minor contributions from haematite.

Because a.f. demagnetization successfully revealed the characteristic remanence in most cases, all specimens were demagnetized at 200 Oe. Subsequently, the reliability of the mean magnetic orientation at each site was statistically evaluated¹⁶ and classified¹⁷ as either 'Class I' (Fisher $k \geq 10$), 'Class II' ($k < 10$, but two samples in close agreement), or 'Class III' (unreliable). All specimens from 'Class II' sites and all normally polarized specimens below the Brunhes/Matuyama boundary were subjected to stepwise thermal demagnetization. This resulted in

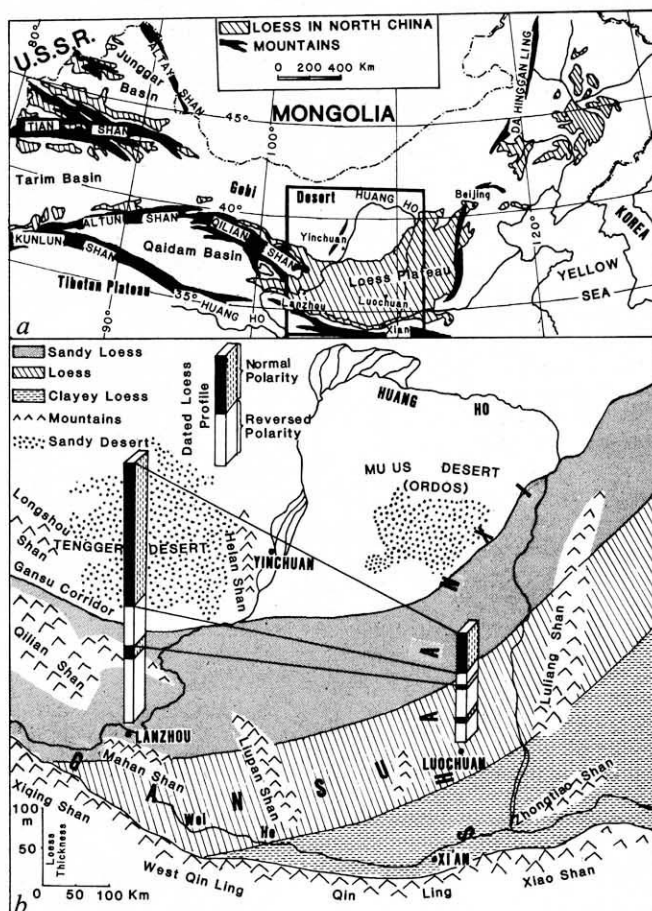


Fig. 2 Lithologies and the magnetic polarity stratigraphy (MPS) from the Lanzhou loess sequence. Only the lower 100 m of the succession are completely exposed. The number of palaeosols depicted in the upper 230 m reflects those occurring near sampling sites and probably underestimates their true abundance. Sampling site numbers (referred to in the text) are shown adjacent to the lithological column. The latitude (degrees) of the virtual geomagnetic pole (VGP) is based on the mean inclination and declination of each site. An α_{95} confidence envelope is plotted (shading) around each VGP latitude. As discussed in the text, sites 16–18 are not included in the VGP or MPS plots. The ages (Myr) of the magnetic boundaries are based on refs 17 and 18. The age of the basal loess is estimated by extrapolation of the mean loess-accumulation rate.

The resulting magnetostratigraphy for the Lanzhou loess is shown in Fig. 2. It comprises 34 Class I sites and 6 Class II sites. Three Class I sites (16–18, Fig. 2) are not included in the magnetic zonation and are discussed below. The latitude of the virtual geomagnetic pole (VGP) is used to define the polarity of each site. The α_{95} confidence envelope plotted for each VGP latitude (Fig. 2) indicates a high level of reliability for most sites. The uppermost normal magnetozone is interpreted as the Brunhes chron, which overlies reversed magnetozone attributed to the Matuyama chron. The Jaramillo normal subchron is defined by sites 13 and 14. Calculations of mean loess-accumulation rates between reversal boundaries of known ages^{18,19} consistently yield a value averaging 26 cm kyr⁻¹ for the past million years. Extrapolation of this rate to the base of the Lanzhou loess sequence suggests that it dates from ~1.3 Myr ago. These results differ from the earlier magnetostratigraphical studies of Wang¹⁵ in that the Brunhes/Matuyama boundary is placed 50 m higher, the base of the sequence is estimated to be 200–300 kyr older, and the reliability of the sites (as defined by the α_{95} -error

In an effort to provide data on an inaccessible portion of our main section (105–125 m, Fig. 2), three sites were collected from scarps ringing collapse pits 100 m west of the primary sampling gully. These normally magnetized sites (16–18, Fig. 2) were not included in the magnetostratigraphy for the following reasons: (1) their stratigraphical position is uncertain due to covered strata obscuring possible unconformities^{12,13}; (2) they may have slumped from Brunhes-aged strata above; and (3) their inclusion would render the magnetostratigraphy less interpretable. Furthermore, our laboratory experiments indicate that periodic wetting to near saturation can cause the magnetic orientation of loess specimens to shift dramatically over periods of less than a week towards the ambient field direction. The amount and stability of the shift appears to be an inverse function of the degree of lithification. The older, well lithified specimens appear

largely unaffected by wetting. In contrast, the major magnetic contributors to the remanence of the younger specimens, including some from sites 16–18, rotate partially or completely into the ambient field direction and are directionally stable against at least a 1,000-Oe demagnetizing field. In natural situations, these experimental results imply that gradual lithification and removal from the zone of periodic wetting by burial are both important components in 'locking in' the characteristic remanence. Because the partially indurated specimens from sites 16–18 lie along collapse pits that were formed by and presently act as conduits for runoff, there is a strong possibility that their characteristic remanence has been reoriented. Consequently, these sites are excluded from the magnetostratigraphy. The specimens from the overlying sites were collected at least 2 m below the modern surface, where they were below the zone of natural wetting and, hence, are unlikely to have been reoriented.

A comparison of the results from Lanzhou with those from Luochuan (Fig. 1b) yields some new insights into the nature of loess deposition in China. Large, modern dustfalls tend to originate in the northern deserts and to distribute loess across the entire loess plateau²⁰. Similarly, ancient loess deposition beginning in the late Pliocene^{4,5} is likely to have been generally synchronous across the loess plateau. Our study indicates, however, that the oldest preserved loess can vary in age between regions by as much as 1 Myr. We attribute these differences to Plio-Pleistocene uplift of the Tibetan Plateau^{11,21} which caused structural disruption of its marginal regions, including those around Lanzhou. Although the ongoing Indo/Asian collision continues to deform north China^{22,23}, major deformation appears to have been transferred out of the Lanzhou basin after the early Pleistocene.

Whereas the loess-accumulation rates have averaged 7 cm kyr⁻¹ at Luochuan since the Jaramillo subchron, the mean rate at Lanzhou has been nearly four times higher during the same interval. This high rate and the remarkable thickness of the Lanzhou loess is explained both by its proximity to the northern deserts and by the local topography, whereby dust-laden, high-velocity winds funnelled through the Gansu Corridor (Fig. 1b) and over Wushaoling Pass expand, decelerate, and release some of their sediment load in the Lanzhou basin. The coarser mean grain size of the Lanzhou loess, when compared to the silty (Luochuan) or clay-rich (Xian) loess farther south (Fig. 1b), reflects similar causes. Palaeomagnetic studies at Luochuan^{4,5} indicate that haematite is the primary magnetic carrier and that it results from *in situ* chemical processes related to lesser or greater amounts of soil formation. In Lanzhou, the high rates of loess accumulation generally overwhelmed the soil-forming process. Consequently, detrital magnetite or titanomagnetite, rather than chemical haematite, is the primary magnetic mineral. Finally, within the completely exposed, lower 100 m of the Lanzhou sequence, 15 weakly to moderately developed palaeosols are preserved (Fig. 2). Given the calculated accumulation rates, one soil-forming episode would have occurred approximately every 25 kyr. This frequency is very similar to that determined for the precession of the equinoxes²⁴ and may, thus, reflect control exerted by Milankovitch-type orbital parameters²⁵ on the early Pleistocene climates of north China.

This study was supported by the Lanzhou University and the Shell Foundation provided support to Douglas W. Burbank. During the experimental studies, Li Jijun was supported as a visiting scholar to the University of Southern California by Lanzhou University. Helpful discussions with S. Lund and a review by F. Theyer are gratefully acknowledged.

Received 19 March; accepted 10 June 1985.

1. Liu, T. S. & Chang, T. H. *Rep. 6th int. Congr. Quat. Warsaw* 6, 503–524.
2. Liu, T. S., Wang, T. M., Wang, K. L. & Wen, L. C. *Sci. Rec.* 2, 167–174 (1958).
3. Sasajima, S. & Wang, Y. Y. *The Recent Research of Loess in China* (Kyoto University and Northwest University, 1984).
4. Heller, F. & Liu, T. S. *Nature* 300, 431–433 (1982).
5. Heller, F. & Liu, T. S. *Geophys. J. R. astr. Soc.* 77, 125–141 (1984).
6. Chen, D. N. *et al. Proc. 3rd Nat. Quat. Congr., China* (1979).
7. Liu, T. S., Wen, O. Z., An, Z. S. & Zheng, H. H. *Int. geol. Congr., Paris* (1980).
8. Zheng, H. H., in *Quaternary Geology and Environment of China* (ed. Liu, T. S.) 59–66 (China Ocean Press, Beijing, 1982).

9. Woo, R. K. *Kexue Tongbao* 6, (1966).
10. Wang, Y. Y. & Xue, S. S. in *Loess and Quaternary Geology* (ed. Wang, Y. Y.) 88–98 (People's Press, Shaanxi Province, Xian, 1982).
11. Li, J. *et al. Scient. Sinica* 22, 1314–1328 (1979).
12. Kukla, G. J. in *After the Australopithecines* (eds Butzer, K. W. & Isaac, G. L.) 99–188 (Mouton, The Hague, 1975).
13. Fink, J. & Kukla, G. J. *Quat. Res.* 7, 363–371 (1977).
14. Shackleton, N. J. *et al. Nature* 307, 620–623 (1984).
15. Wang, Y. Y., in *Loess and Quaternary Geology* (ed. Wang, Y. Y.) 20–47 (People's Press, Shaanxi, 1982).
16. Fisher, R. A. *Proc. R. Soc. A* 217, 295–305 (1953).
17. Johnson, N. M. *et al. Palaeogeogr., Palaeoclimatol., Palaeoecol.* 37, 17–42 (1982).
18. Mankinen, E. A. & Dalrymple, G. B. *J. geophys. Res.* 84, 615–626 (1979).
19. Harland, W. B. *et al. A Geologic Time Scale* (Cambridge University Press, 1982).
20. Liu, T. S., Gu, X. F., An, Z. S. & Fang, Y. X. in *Desert Dust: origin Characteristics and Effect on Man* (ed. Pewe, T. L.) 149–158 (Geological Society of America, Boulder, 1981).
21. Xu, R. in *Geological and Ecological Studies of the Qinghai-Xizang Plateau Vol. 1* 139–144 (Science Press, Beijing, 1981).
22. Tapponnier, P. & Molnar, P. *J. geophys. Res.* 82, 2905–2930 (1977).
23. Molnar, P. & Tapponnier, P. *J. geophys. Res.* 83, 5361–5375 (1978).
24. Berger, A. in *Milankovitch and Climate, Part I* (eds Berger, A., Imbrie, J., Hays, J., Kukla, G. & Saltzman, B.) 3–39 (Reidel, Dordrecht, 1984).
25. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* 214, 1121–1132 (1976).

Effect of drought on dust production in the Sahel

N. J. Middleton

School of Geography, University of Oxford, Mansfield Road, Oxford OX1 3TB, UK

The severe drought currently afflicting the Sudano-Sahelian zone to the south of the Sahara Desert has been suggested to be instrumental in producing an increased output of soil-derived aerosols into the atmosphere from the region¹. During the very dry period 1972–74 mean aerosol concentrations at Barbados, West Indies, as affected by the African Dust Plume², were three times that of pre-drought levels¹, that is before 1968. A marked increase in the frequency of severe dust occurrences in northern Nigeria has also been noted during 1972 and 1973 (ref. 3). I present here data from selected meteorological stations which show that dust-storm activity in the west and east of the Sudano-Sahelian belt has dramatically increased during the drought years; by a factor of 6 in Mauritania and up to a factor of 5 in Sudan.

The Sudano-Sahelian zone lies approximately between latitudes 10 and 20° N and includes parts of Mauritania, Senegal, Mali, Burkina Faso, Niger, Chad, Sudan and Ethiopia. Study of the zone is impaired by a general lack of adequate long-term data for large parts of its area, but data for Mauritania in the west and Sudan in the east have made this investigation possible.

Perhaps the most striking representation of an increase in dust-raising activity with the onset of drought conditions is shown in Fig. 1. This shows the variation of annual dust-storm frequency and annual rainfall totals for Nouakchott in Mauritania and El Fasher in Sudan. At Nouakchott, and generally for Mauritanian stations south of ~20° N, dust-storm activity is largely concentrated in the period from January to May, before the onset of the rainy season. At this time of year, dust concentrations at Barbados are relatively low¹, but the dust concentration measured at Cayene, French Guiana⁴ is at a maximum. (North of 20° N, where the rainy season starts later in the year, maximum dust-storm activity starts in February and continues until July or August.) Low rainfall totals of 48.1 mm in 1970 and 17.9 mm in 1971 represented just 32 and 12%, respectively, of the 1949–67 average, and can be seen as the main onset of the drought. The number of dust-storm days increased dramatically from 6 in 1970 to 65 in 1974 before a reasonably high annual rainfall of 190.6 mm in 1975; dust-storm activity declined to 25 days in 1976 and 27 days in 1977. In 1977, the rainy season brought only 2.7 mm of precipitation, however, making it the driest year since records began in 1931, and dust-storm activity rose to 55 and 61 days in 1978 and 1979 respectively. The total dropped to 33 dust-storm days in 1980 after a relatively heavy rainfall in 1979, but rose to an unprecedented 85 days in 1983 as rainfall again diminished.

A similar pattern is shown in the graph for El Fasher in the eastern Sahel (Fig. 1b). Here particularly low rainfall in 1972 and 1973 was followed by a distinct rise in dust-storm frequency

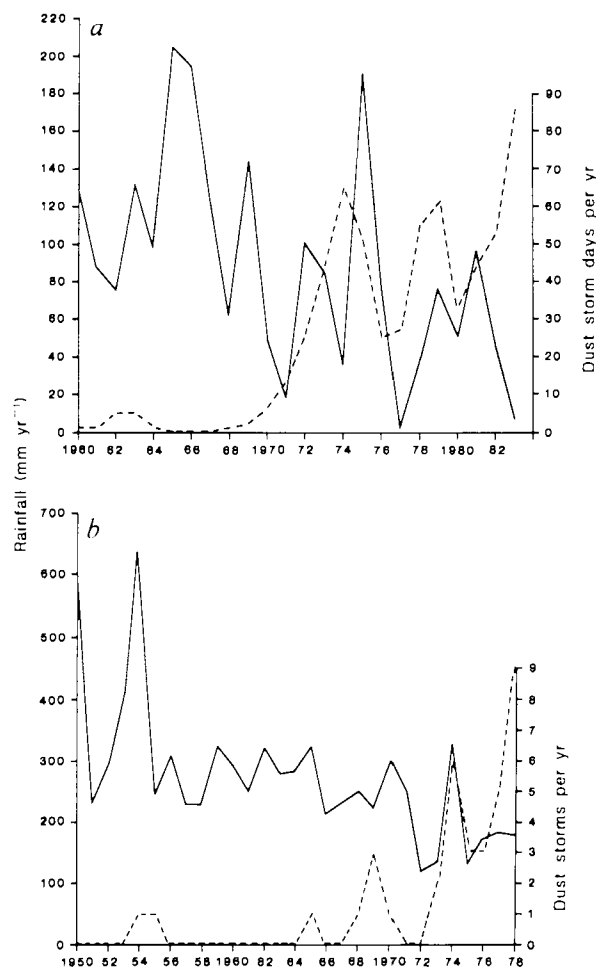


Fig. 1 Annual dust storm frequencies (dotted line) and annual rainfall (solid line) totals at *a*, Nouakchott (18°07' N, 15°56' W) and *b*, El Fasher (13°32' N, 25°21' E). Dust-storm data for Nouakchott are in the form of 'dust-storm days', that is, a day with a dust storm occurring (data from the Service Météorologique, Nouakchott); mean annual number = 25.3. Dust-storm data for El Fasher are the number of dust storms occurring (data from Sudan Meteorological Department Annual Meteorological Reports); mean annual number = 1.2. A dust storm is defined as an event reducing visibility to <1,000 m.

peaking at six storms in 1974, falling in 1975 and 1976 after high rainfall in 1974, and rising once more to nine per year in 1978 as rainfall remained below 58% of the 1950–67 average from 1975 to 1978. Similar trends were found for other stations in the Sudanese Sahel from which data were available: El Obeid, Khartoum, Tokar and Aroma.

Although Fig. 1 illustrates an increase in dust-storm activity dating from the early 1970s at Nouakchott and El Fasher, the quality of the data used does not unfortunately give any indication of the duration of dust-raising events, and thus prevents a quantitative assessment of the magnitude of change in the quantities of material raised. This problem is resolved for Nouakchott in Fig. 2, which shows the average monthly percentage frequency of dust storms by hour for two periods: 1949–67, representing the pre-drought period, and 1968–77, covering the first 10 yr of the drought. The differences in frequencies between the two periods are marked. The frequency during the peak time of activity, 1200 to 1500 LT, in the drought years being on average nearly seven times that of the pre-drought period during the dustiest months of January to April. The increase in terms of numbers of dust-storm events between the two periods is smaller, being just over three times more frequent during drought years. Similar comparative data were available for seven other stations in the Mauritanian Sahel; Table 1 summarizes these data to

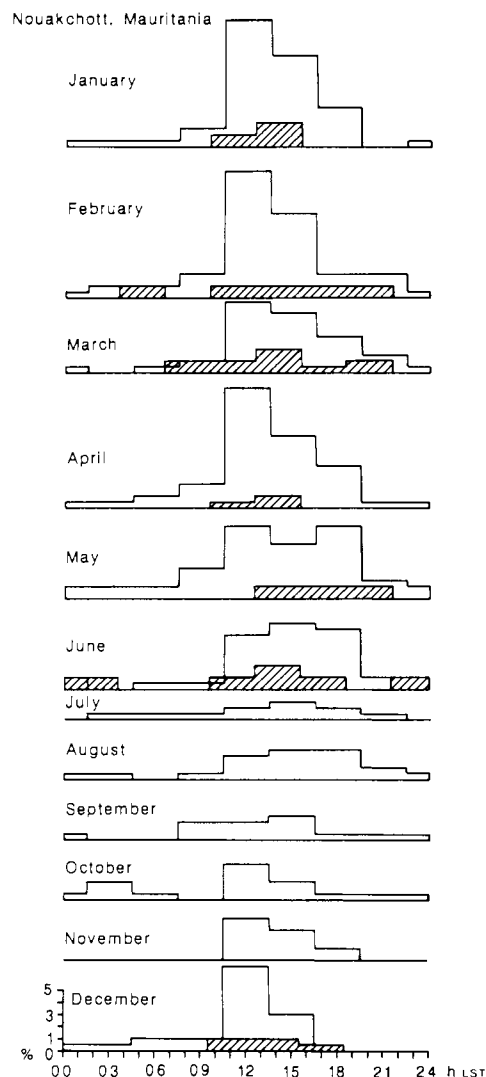


Fig. 2 Diurnal arithmetic mean frequency per cent by month of visibility reduced to <1,000 m resulting from dust storms at Nouakchott. The hatched area represents frequency per cent for the period 1949–67 (ref. 5); the unhatched area represents frequency per cent for the period 1968–77 (data from the Service Météorologique, Nouakchott).

show the increase in average diurnal frequency per cent from the pre-drought period to the drought years. Combining these results shows that, on average, dust-raising activity in the Mauritanian Sahel was more than six times greater during the drought as compared with the pre-drought years.

Coherent rainfall data covering these periods were available only for Nouakchott, which indicates that the average for the 1968–77 period was 51% of the pre-drought average. However, recent reports on sub-Saharan rainfall^{6,7} that have included data from some of these Mauritanian stations, indicate that the drought which started in 1968 has persisted strongly throughout the 1970s, and, indeed, still continues. It seems clear, therefore, that the quantity of material raised by wind has greatly increased during a period of substantially reduced rainfall in the Sahel of Mauritania.

Further analysis of the Sudanese Sahel was carried out using the relatively gross variable of mean numbers of dust storms occurring during pre-drought and drought years (Table 2). Here the averages for 1950–67 and 1968–78 are shown with the increase between the two periods, and corresponding decrease for rainfall. No mean increase in dust storms is shown for all five stations as this would lead to biasing by the stations that had few storms previously and now experience many more storms, but still relatively few compared with other stations

Table 1 Diurnal arithmetic frequency per cent and increase of dust storms at stations in the Mauritanian Sahel for two periods: 1949–67; 1968–77

Hour (LST)	Nouakchott			Kiffa			Nouadhibou			Rosso		
	Frequency %	Frequency %	Increase*	Frequency %	Frequency %	Increase*	Frequency %	Frequency %	Increase*	Frequency %	Frequency %	Increase*
	1949–67	1968–77		1949–67	1968–77		1949–67	1968–77		1949–67	1968–77	
00	0.1	0.5	500	0.2	1.2	600	0.1	1.3	1,300	0.1	0.0	—
03	0.1	0.4	400	0.0	0.8	—	0.2	1.4	700	0.0	0.0	0
06	0.1	0.6	600	0.1	0.6	600	0.3	2.6	867	0.0	0.0	0
09	0.4	1.0	250	0.6	1.0	167	0.3	2.0	667	0.0	0.1	—
12	0.8	5.4	675	0.2	1.7	850	0.9	0.9	0	0.1	0.9	900
15	0.3	4.1	1,367	0.1	1.6	1,600	0.9	1.3	144	0.2	0.7	350
18	0.3	2.4	800	0.3	1.6	533	0.5	2.3	460	0.1	0.7	700
21	0.1	0.7	700	0.0	1.0	—	0.3	0.7	233	0.0	0.0	0
			662			725			546			325
	Tidjika			Akjoujt			Atar			Boutilimit		
	Frequency %	Frequency %	Increase*	Frequency %	Frequency %	Increase*	Frequency %	Frequency %	Increase*	Frequency %	Frequency %	Increase*
	1949–67	1968–77		1949–67	1968–77		1949–67	1968–77		1949–67	1968–77	
00	0.0	3.6	—	0.0	3.6	—	0.5	1.5	300	0.0	1.7	—
03	0.0	0.6	—	0.0	0.8	—	0.6	1.4	233	0.1	0.8	800
06	0.1	0.8	800	0.3	0.8	267	0.5	1.5	300	0.3	1.4	467
09	0.2	1.2	600	0.5	1.9	380	0.6	1.6	267	0.4	1.7	425
12	0.2	2.0	1,000	0.3	2.4	800	0.3	2.7	900	0.4	3.9	975
15	0.1	1.4	1,400	0.2	1.4	700	0.3	1.8	600	0.3	3.2	1,067
18	0.3	0.8	267	0.0	0.7	—	0.3	1.6	533	0.1	2.3	1,300
21	0.0	1.7	—	0.0	3.4	—	0.4	1.2	300	0.0	3.4	—
			813			537			429			1,005

The mean increase between 1949–67 and 1968–71 is 630. Data for 1949–67 is from ref. 5; data for 1968–77 from the Service Meteorologique, Nouakchott. * 1968–77 values are expressed as percentage of 1949–67 values.

Table 2 Arithmetic mean number of dust storms and mean rainfall per year at stations in the Sudanese Sahel for the two periods, 1950–67 and 1968–78 and the corresponding increase/decrease

Station	Location	Average dust storms (mm yr ⁻¹)			Average rainfall (mm yr ⁻¹)		
		1950–67	1968–78	Increase*	1950–67	1968–78	Decrease*
El Fasher	13°32' N 25°21' E	0.2	3.0	1,500	317.5	208.5	66
El Obeid	13°11' N 30°14' E	0.7	3.6	514	400.1	301.6	75
Khartoum	15°33' N 32°35' E	1.8	8.6	478	186.4	134.7	72
Aroma	15°50' N 36°09' E	3.3	6.5	197	245.7	189.4	77
Tokar	18°26' N 37°44' E	36.6	47.2	129	90.8	55.5	61
							Mean 70

* 1968–78 values are expressed as percentage of 1950–67 values (data from Sudan Meteorological Department Annual Meteorological Reports).

(such as, El Fasher). Nevertheless, the pattern is clear; for five stations across the Sudanese Sahel the number of dust storms occurring during the drought has markedly increased while rainfall in the region averaged 70% of the pre-drought average.

These data also clearly demonstrate the importance of environmental and terrain effects. Tokar, the driest station, has by far the greatest frequency of dust storms, recording more per year than all the other stations combined even during the height of the drought. It seems likely that this high incidence is due in part to Tokar's location at the end of a long, deep valley, which both concentrates runoff (carrying fine sediments into readily deflated locals) and funnels winds, thereby yielding a higher percentage of days with winds above the critical threshold velocity⁸.

Correlations between long-term rainfall measures and dust-storm frequency are poor⁹; the relationship between drought and dust production is complex. Entrainment of soil material by wind is enhanced if the land surface is composed of fine-grained, unconsolidated material that is sparsely covered by protective vegetation. Given a fine-grained, potentially wind-erodible soil the lack of moisture resulting from drought conditions will directly reduce the likelihood of the soil particles aggregating as larger, less erodible units. Furthermore, the lack of moisture will have an indirect effect on the erodibility of a soil by reducing vegetation cover over a period of some years. These effects can also be exacerbated by the actions of man in several ways, so that the tilling of dry soils helps to break down soil aggregates, thus increasing its wind erodibility; overgrazing of drought-afflicted vegetation can lead to a reduction of this

protective cover, as can the collection of wood for fuel. These natural and anthropogenic factors, often operating in conjunction, form the main elements of the 'desertification' process, and may also increase the probability of soil erosion by wind in certain areas. Indeed, land use impacts may be especially relevant in this study as all the meteorological stations are in population centres. Thus, agricultural activities and livestock grazing may contribute to an increased dust mobilization locally when compared with the region as a whole.

The increase in dust-storm activity noted above could be attributable to other meteorological factors, such as increased wind speeds, gustiness, turbulence and alterations in circulation patterns, and possibly any or all of these factors may have some bearing on this phenomenon. However, with the knowledge of the effects of a reduced rainfall on wind erodibility of a soil as outlined above it is reasonable to suggest that drought in the Sahel, in conjunction with poor land-use practices that have been noted by several workers^{10,11}, is a major influential factor in the observed increase of dust-storm activity in the region.

There are many environmental implications of this dramatically enhanced aerosol output. Effects on weather and climate are derived from the importance of dust on the radiation budget of the atmosphere, which operate in two ways¹²: directly by changing the fluxes in a cloud-free atmosphere by scattering and absorption of radiation; and indirectly by modifying the optical properties, particularly the albedo of clouds. Perhaps the most significant meteorological effect of these influences for the Sahel can be seen as their enhancement of stability in the planetary boundary layer. The base of the Saharan air layer is

defined by a persistence of a temperature inversion that is still recognizable on meteorological soundings from the Caribbean and in Miami². The dust plume over northern Nigeria is also often associated with a pronounced temperature inversion^{3,13}, and although this nighttime phenomenon is a climatological feature resulting from radiation cooling from the ground surface, the presence of dust in the atmosphere may reinforce the feature and feedback on surface conditions. Such stability has important consequences by its inhibition of convection and potential rain-making processes over the Sahel. Reduced precipitation may also result from the effect of a reduction in atmospheric loads of biogenic ice nuclei resulting from overgrazing during the drought¹⁴.

Recent investigations into feedback mechanisms that may exacerbate drought conditions in the Sahel have focused on ground-surface albedo changes due to overgrazing¹⁵⁻¹⁷ and soil-moisture properties¹⁸. The role of dust in major climatic modification over northwestern India has been investigated¹⁹ and its possible effects over the Sahel should be studied further²⁰ in the light of the dramatic increase in dust-storm activity that has been occurring since the start of the drought.

I thank the meteorological services of Mauritania and Sudan

for their provision of information and data, and Professors A. S. Goudie and J. M. Prospero for useful comments.

Received 18 March; accepted 5 June 1985.

1. Prospero, J. M. & Nees, R. T. *Science* **196**, 1196-1198 (1977).
2. Carlson, T. N. & Prospero, J. M. *J. appl. Met.* **11**, 283-297 (1972).
3. Kalu, A. E. thesis, Univ. Birmingham (1982).
4. Prospero, J. M., Glaccum, R. A. & Nees, R. T. *Nature* **289**, 570-572 (1981).
5. Hinds, B. D. & Hoidale, G. B. *Research and Development Technical Rep. ECOM-DR-77-3* (US Army Electronics Command, White Sands Missile Range, New Mexico, 1977).
6. Lamb, P. J. *Nature* **299**, 46-48 (1982).
7. Lamb, P. J. *J. Climatol.* **3**, 419-422 (1983).
8. Gillette, D. A., Adams, J., Endo, A. & Smith, D. J. *geophys. Res.* **85**, 5621-5630 (1980).
9. Goudie, A. S. *Prog. phys. Geogr.* **7**, 502-530 (1983).
10. Rapp, A. *A Review of Desertization in Africa* (Secretariat for International Ecology, Stockholm, 1974).
11. Ormerod, W. E. *J. Arid Envir.* **1**, 357-379 (1978).
12. Junge, C. in *Saharan Dust: Mobilization, Transport, Deposition*, (ed. Morales, C.) 49-60 (SCOPE Rep. 14 Wiley, New York, 1979).
13. Kalu, A. E. in *Saharan Dust: Mobilization, Transport, Deposition* (ed. Morales, C.) 95-118 (SCOPE Rep. 14 Wiley, New York, 1979).
14. Schnell, R. C. & Brine, C. J. *EOS* **56**, 994-995 (1975).
15. Charney, J. Q. *Jl R. met. Soc.* **101**, 193-202 (1975).
16. Idso, S. B. Q. *Jl R. met. Soc.* **103**, 369-370 (1977).
17. Wendler, G. & Eaton, F. *Climat. Change* **5**, 365-380 (1983).
18. Walker, J. M. & Rowntree, P. R. Q. *Jl R. met. Soc.* **103**, 29-46 (1977).
19. Bryson, R. A. & Baeris, D. A. *Bull. Am. met. Soc.* **48**, 136-142 (1967).
20. Bryson, R. A. *Science* **184**, 753-760 (1974).

Extreme carbon-isotope enrichments in evaporating brines

M. Stiller, J. S. Rounick & S. Shasha

Isotope Department, Weizmann Institute of Science, Rehovot 76100, Israel

We report here extreme ¹³C enrichments in the dissolved inorganic carbon pool (ΣCO₂) in evaporating brines. These enrichments may influence the interpretation of the stable carbon isotope record of both inorganic and organic carbon deposited in saline environments. Carbon isotope values of ΣCO₂ respond to types of carbonate inputs, rate of exchange with the atmospheric CO₂ and to biological effects^{1,2}. δ¹³C values of up to +14‰ (all values are reported here with respect to the PDB standard) for ΣCO₂ from anoxic waters have been reported where methane production by CO₂ reducing bacteria occurs³. We document an abiotic effect in evaporating brines which resulted in ¹³C enrichments of up to +16.5‰ in natural conditions (solar evaporation ponds used for potash production) and up to +34.9‰ in laboratory evaporation experiments, the most enriched δ¹³C values reported to date in oxic conditions. These extreme enrichments may result from a non-equilibrium gas-transfer isotope fractionation.

The production of potash from the Dead Sea involves natural solar evaporation in which the Dead Sea brine (total dissolved solids (TDS) ~340 g l⁻¹) is concentrated in a series of steps. First, NaCl is precipitated out by this process, and when the solution reaches a TDS of ~440 g l⁻¹, carnallite (KCl·MgCl₂·6H₂O) begins to precipitate. Finally, when the brines reach a TDS of 494 g l⁻¹, these 'end brines' are discharged back into the Dead Sea after their precipitates have been collected for industrial use⁴.

Physical and chemical characteristics were determined for these evaporating brines at various stages of concentration. Brines were sampled on 17 July 1984; temperature and pH were measured on site. ΣCO₂ and δ¹³C were sampled by filling, on site, evacuated vessels containing concentrated H₃PO₄ under vacuum and then measured by extracting the evolved CO₂ on a vacuum system in the laboratory. The volume of ΣCO₂ was determined manometrically and its δ¹³C (measured with respect to the international PDB standard) determined on a MAT Varian 250 ratio mass spectrometer. Samples for methane and oxygen determinations were also collected on site. Oxygen was determined using the modified Winkler method⁵ and methane deter-

mined by gas chromatography (J.S.R. and A. Serban, unpublished data).

The evaporating brines showed the highest ¹³C enrichments ever reported for a water body open to the atmosphere (Table 1). These extreme ¹³C enrichments were also found in two samples of end brine collected in 1980 and 1983. Their δ¹³C values were +16.8‰ and +16.0‰, respectively; none of the brines had detectable methane and all were oxygenated (~1 mg l⁻¹). Therefore, these ¹³C enrichments do not appear to be associated with CO₂ reduction by methanogenic bacteria. ¹³C enrichment was inversely related to ΣCO₂ and as evaporation proceeds, an increase in density (salinity) occurs with a concurrent decrease in pH and ΣCO₂ content (Table 1). This suggests that evaporation effects related to the loss of CO₂ from the brines are causing the ¹³C enrichments.

Evaporation proceeds in the ponds in variable environmental conditions. Thus, the ¹³C enrichments and related ΣCO₂ decrease from pond to pond do not necessarily represent continuous evaporation in constant conditions. Therefore, we conducted a laboratory experiment to control and also to extend the evaporation range. Dead Sea water sampled on 3 July 1984 and brine from Evaporation Pond 1 (EP1) were evaporated at 45°C in quiescent conditions for 8 and 41 days, respectively.

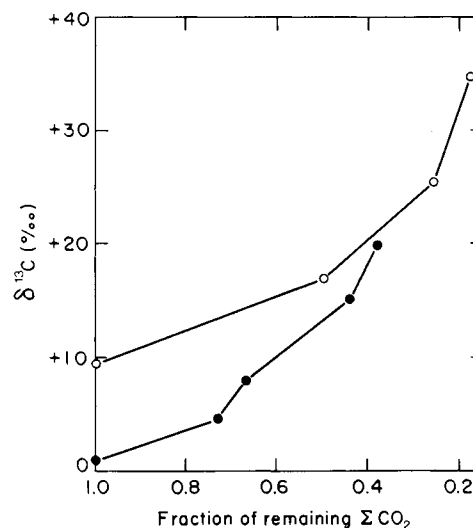


Fig. 1 Relationship between 'normalized' remaining fraction of ΣCO₂ and δ¹³C(ΣCO₂) for Dead Sea brine (●) and Evaporation Pond 1 (○) experiments.

Table 1 Sample description, total CO₂(Σ CO₂) and $\delta^{13}\text{C}$ of evaporating brines for four solar evaporation ponds (EP1-EP4)

Site*	Temperature (°C)	Density† (g cm ⁻³)	pH	Σ CO ₂ (mmol per kg brine)	$\delta^{13}\text{C}$ (‰)
EP1	34.5	1.2835	5.3	0.61	+10.4
EP2	36.0	1.2923	5.2	0.76	+7.9
EP2	37.1	1.3103	4.9	0.59	+13.3
EP4	40.3	1.3373	4.5	0.58	+16.5

* All EP samples collected on 17 July 1984.

† Density measured pycnometrically at 40 °C and calculated at sampling temperature with a thermal expansion coefficient of $4.2 \times 10^{-4} \text{ g cm}^{-3} \text{ °C}^{-1}$ (ref. 13).

Samples were collected periodically during the experiment and analysed for pH, density, Σ CO₂ and $\delta^{13}\text{C}$ (Σ CO₂) (Table 2). Ca + Mg molarities were measured by EDTA titrations and were regarded as conservative indicators of brine concentration during evaporation.

There is a dramatic increase in $\delta^{13}\text{C}$ (Σ CO₂) in both experiments, up to +19.8‰ in the evaporated Dead Sea water and +34.9‰ in the evaporated brine of pond EP1 (Table 2). These enrichments occur concurrently with an increase in acidity of the brines and resultant Σ CO₂ losses. The remaining fraction of Σ CO₂ during the experiment was estimated by normalizing the Σ CO₂ at each sampling step to the increasing concentration during evaporation of Ca + Mg. (The behaviour of Ca + Mg is not completely that of conservative species because some Mg losses occur during the precipitation of carnallite and Ca losses with precipitation of gypsum and/or calcium carbonate.) We calculate the fraction of remaining Σ CO₂ in the final samples of the Dead Sea brine and EP1 experiments to be 0.38 and 0.18 respectively. The fraction of remaining Σ CO₂ plotted against the $\delta^{13}\text{C}$ (Σ CO₂) suggest a Rayleigh distillation type enrichment (Fig. 1). We surmise that the decrease in Σ CO₂ is due to CO₂ escape from the brine because of increasing salinity and acidity as evaporation proceeds. By fitting the data points shown in Fig. 1 to a Rayleigh type curve, we obtain apparent ϵ values of -19.4 and -14‰ for the Dead Sea and EP1 experiments respectively. In $\delta(\%)$ nomenclature, the Rayleigh equation can be expressed as $\delta = \delta_0 + \epsilon \ln f$, where δ_0 , δ are the initial and final $\delta^{13}\text{C}$ (‰) values of the Σ CO₂ pool respectively, $\epsilon(\%) = (\alpha - 1)1,000$, α is the isotopic fractionation factor, f is the fraction of Σ CO₂ remaining in solution⁶. These 'apparent' ϵ s are much larger than the equilibrium ϵ between gaseous CO₂ and HCO₃⁻ at 45 °C which is -7.5‰ (ref. 7).

One possible explanation for this discrepancy is that in the above calculation we have calculated the remaining fraction of Σ CO₂ and refer to it as that of HCO₃⁻. Unfortunately, bicarbonate cannot be independently determined for Dead Sea brines because of high boron concentrations⁸, 45 mg l⁻¹, which interfere with alkalinity titrations. However, Saas and Ben Yaakov⁹ suggest that an increase in brine salinity and concurrent pH decrease does not necessarily imply a higher pCO₂, but, in fact, an enhancement of CaCO₃ solubility and Mg bicarbonate complexing. Thus, by assuming that most of the Σ CO₂ is in bicarbonate form, as implied by the small fraction of 'free CO₂' evolved under vacuum (it ranged from 10 to 25% of the Σ CO₂ between the initial and final stages of evaporation experiments; our unpublished data), we can discount this explanation.

We are therefore left with the contention that loss of CO₂ occurs during the evaporation process in non-equilibrium conditions. Kinetic fractionation will thus affect ϵ , making it larger than its equilibrium value. The theoretical maximum single-stage kinetic fractionation for loss of gaseous CO₂ is given by $\alpha = (44/45)^{1/2}$, which is equivalent to -11.2‰ and with a value of -7.5‰ for equilibrium fractionation at 45 °C (the difference between $\delta^{13}\text{C}$ values of the 'free CO₂' and 'bicarbonate' at the various stages of our experiments remained close to this expected

Table 2 Sample description, total CO₂ and $\delta^{13}\text{C}$ in brines evaporated at 45 °C; laboratory experiments

Step	Density at 45 °C (g cm ⁻³)	pH	Ca + Mg (mol per kg brine)	Σ CO ₂ (mmol per kg brine)	$\delta^{13}\text{C}$ (‰)
Dead Sea brine					
Initial*	1.226	6.0	1.86	0.75	+0.91
DS-1	1.241	5.6	2.10	0.62	+4.5
DS-2	1.256	5.5	2.36	0.64	+7.8
DS-3	1.285	5.1	2.88	0.51	+15.0
DS-4	1.309	4.9	3.22	0.50	+19.8
Evaporation Pond 1 brine					
Initial	1.279	5.3	2.83	0.54	+9.6
EP1-1	1.322	4.7	3.43	0.33	+16.8
EP1-2	1.356	4.2	3.88	0.19	+25.2
EP1-3	1.365	3.7	4.12	0.14†	+34.9

* Dead Sea brine collected on 3 July 1984, 0.5 m depth.

† Estimated after cleaning from the HCl which evolves from the brine upon acidification.

equilibrium value; our unpublished data), the total fractionation would be -18.7‰. The apparent ϵ of -19.4‰ calculated for the Dead Sea brine experiment thus suggests a two-stage process in which there is rapid equilibrium maintained between HCO₃⁻ and dissolved CO₂ but no equilibrium between escaping gaseous CO₂ and dissolved CO₂.

Note that the 'apparent' ϵ for the EP1 experiment is closer to the equilibrium ϵ than that of the Dead Sea experiment. This is quite plausible because the EP1 experiment proceeded at a much slower rate than the Dead Sea experiment. At the very high salinities of the EP1 experiment, the saturation vapour pressure approaches the ambient humidity (as the water activity in the more saline brines is decreasing, the relative humidity approaches unity¹⁰). Therefore, the rate of evaporation slows down and, as a result, the CO₂ escape rate decreases and conditions become more favourable to approach equilibrium conditions of fractionation.

The $\delta^{13}\text{C}$ enrichments of the Σ CO₂ pool in evaporating brines may be recorded sedimentologically by both inorganic and organic processes occurring in these brines. First, our data suggest that $\delta^{13}\text{C}$ values of carbonates formed in association with evaporating brines should be $\delta^{13}\text{C}$ enriched, reflecting the $\delta^{13}\text{C}$ enrichment of the CO₂ pool. Such enriched carbonates were indeed found by Katz *et al.*¹¹ in the aragonite laminae in deposits from the saline Lake Lisan, the Pleistocene precursor of the Dead Sea; they noted that increasing Na/Ca ratios (an indicator of salinity) were positively correlated with $\delta^{13}\text{C}$ increases in these deposits. Second, $\delta^{13}\text{C}$ enrichments in Σ CO₂ of evaporating brines may be responsible for occurrences of heavy $\delta^{13}\text{C}$ values in organic matter deposits from saline environments and may explain the 'superheavy' $\delta^{13}\text{C}$ values of photosynthetic microbial mats found in coastal brine pools in the Sinai Peninsula¹².

We thank Professor J. Gat for valuable advice and for reviewing the manuscript. J. Harash of the Dead Sea Works Ltd guided us in sampling the evaporation ponds and contributed useful background information.

Received 13 March; accepted 5 June 1985.

- Turner, J. V., Fritz, P., Karrow, P. F. & Warner, B. G. *Can. J. Earth Sci.* **20**, 599-613 (1983).
- Rounick, J. S. & James, M. R. *Limnol. Oceanogr.* **29**, 386-389 (1984).
- Stiller, M. & Magaritz, M. *Limnol. Oceanogr.* **19**, 849-853 (1974).
- Epstein, J. A. *Hydrometallurgy* **2**, 572-576 (1976).
- Carpenter, J. H. *Limnol. Oceanogr.* **10**, 141-144 (1965).
- Gat, J. in *Stable Isotope Hydrology*, 21-33 (IAEA, Vienna, 1981).
- Mook, W. G., Bommerson, J. C. & Staverman, W. H. *Earth planet. Sci. Lett.* **22**, 169-176 (1974).
- Schonfeld, I. & Held, S. *Isr. Atom. Energy Comm.* **1A-1061** (1965).
- Sass, E. & Ben-Yaacov, S. *Mar. Chem.* **5**, 183-199 (1977).
- Gat, J. R. *Earth planet. Sci. Lett.* **71**, 361-376 (1984).
- Katz, A., Kolodny, Y. & Nissenbaum, A. *Geochim. cosmochim. Acta* **41**, 1609-1626 (1977).
- Schidlowski, M. & Matzigkeit, U. *Naturwissenschaften* **71**, 303-308 (1984).
- Steinhorn, I. *Isr. J. Earth Sci.* **29**, 191-196 (1980).

Palaeolimnological evidence that lake acidification is accompanied by loss of organic matter

Ronald B. Davis, Dennis S. Anderson & Frode Berge*

Department of Botany and Plant Pathology, University of Maine, Orono, Maine 04469, USA

The hypothesis that loss of dissolved organic matter from lake water accompanies acidification^{1,2} has been investigated palaeolimnologically for Hovvatn (pH 4.4; TOC (total organic carbon) 3.2 p.p.m.) and Holmvatn (pH 4.7; TOC 2.2 p.p.m.) in southern Norway. Regression equations^{3,4} that relate sedimentary diatom remains separately to pH and TOC were applied to diatom counts from ²¹⁰Pb-dated cores from the lakes. The results indicate that in pre-industrial times both lakes were acidic (pH 4.9–5.1, no bicarbonate alkalinity) and relatively humic (TOC 6–9 p.p.m.). This acidification started about 1920 at Hovvatn and in the 1940s at Holmvatn, and was accompanied by TOC decreases of 3–6 p.p.m. This supports the hypothesis and suggests that the acidification of such lakes transformed them from organic weak acid dominated to mineral strong acid dominated. It also implies that acidification decreased the availability of organic ligands for binding potentially toxic aluminium.

The understanding of the acidification of lakes through acidic inputs from the atmosphere has been limited by a paucity of direct measurement of pH and related water chemistry before 25 yr ago. Palaeolimnological reconstructions of lake-pH based on diatom remains in sediment has helped to fill this gap^{5–7}. Using this method, we have inferred for several lakes in southern Norway that: (1) acidification began at differing dates between 1850 and 1940; (2) the total decreases in pH have been <0.8 unit; and (3) the acidified lakes that now have pH 4.4–4.7 and have lost their fish populations in recent decades⁸ were naturally quite acidic (pH 4.9–5.2) 200 or more years ago, before the heavy industrialization of western Europe⁷. To help account for the fish mortality that occurred despite the surprisingly small decrease in pH (but see 'Note added in proof' below), we are extending the palaeolimnological approach to water chemistry parameters other than pH (ref. 4).

It has been argued^{9–11} that the acidification of such lakes is accompanied by precipitation of the coloured organic matter in the water by acid-mobilized metals (largely Al), thus explaining the clarity of acidified waters. There is limited historical and experimental evidence to support this hypothesis¹². Further evidence is needed to confirm that the acidification and the clarification of the lakes did, in fact, co-occur.

We have studied 35 soft-water lakes in Norway to determine the contemporary relationships between water chemistry parameters^{13,14} and the diatom remains in deep-water surface-sediment^{3,4}. By principal components and other statistical analyses, we have shown⁴ that the patterns in the diatom data from these lakes could be largely explained by two groups of environmental parameters: (1) pH and related factors including alkalinity and ΣAl; and (2) TOC and related factors including water colour and transparency. Dissolved organic carbon (DOC) data were unavailable, but for these lakes TOC is a good indicator of DOC. In such strongly oligotrophic lakes with relatively undisturbed watersheds^{8,13}, TOC resides largely in the DOC and imparts much of the yellow or brown colour that reduces the transparency of the water. Dissolved water colour (DWC) also serves as an indication of DOC. DWC was correlated with TOC (correlation coefficient $r = 0.75$) and ln transparency ($r = -0.77$) (refs 3, 4). DWC was visually estimated on a five-point scale, 0–4 (ref. 13). This led to reduced precision and smaller r values, and made the palaeolimnological reconstruction of DWC less useful than that of TOC.

Table 1 Regressions of TOC concentration on diatom remains in deep-water surface sediment from 35 lakes in southern Norway

Stepwise multiple regression (BMDP2R³⁴) of 10 diatom clusters, retaining clusters 4 and 7*:

$$\text{TOC (mg l}^{-1}\text{)} = 2.058 + 0.106(\% \text{CLUS 4}) - 0.03(\% \text{CLUS 7})$$

$$r^2 = 0.59 \quad F\text{-ratio} = 22.9 \quad S_e = 1.08 \text{ mg l}^{-1} \\ (P = 0.001)$$

Stepwise multiple regression of principal components (BMDP4R³⁴) of six taxa†, retaining principal components 1 and 3:

$$\begin{aligned} \text{TOC (mg l}^{-1}\text{)} = & 1.87 - 0.31(\text{ACAUSvHE}) \\ & - 0.051(\text{ACMARGIN}) \\ & - 0.002(\text{OTHERAC}) \\ & + 0.154(\text{ANSERVBR}) \\ & + 0.122(\text{FRUSTULI}) \\ & - 0.011(\text{NAKRASSK}) \end{aligned}$$

$$r^2 = 0.61 \quad F\text{-ratio} = 24.6 \quad S_e = 1.05 \text{ mg l}^{-1} \\ (P = 0.001)$$

* The diatom taxa comprising these clusters are given in ref. 4.

† These are the six taxa most closely related to TOC, but not related to pH (ref. 4): *Achnanthes austriaca* v. *helvetica*, *A. marginulata*, other *Achnanthes* spp., *Anomoeoneis serians* v. *brachysira*, *Frustulia rhomboides* and *Navicula krasskei*.

These analyses⁴ indicate that the diatom data could be used to develop the multiple regression equations¹⁵ which are useful for separate reconstructions of past pH and TOC. The pH equations and errors have been given elsewhere³. Here Table 1 lists the TOC equations and their errors. The regression standard errors (S_e) were used to estimate the errors on down-core pH and TOC inferences, as in ref. 16, but we have pointed out³ that much of the variance in the calibrational (35 lake) data set is not relevant to a series of inferences in a core from a single lake. The changes that we infer down-core are probably more accurate than the regression standard errors would suggest.

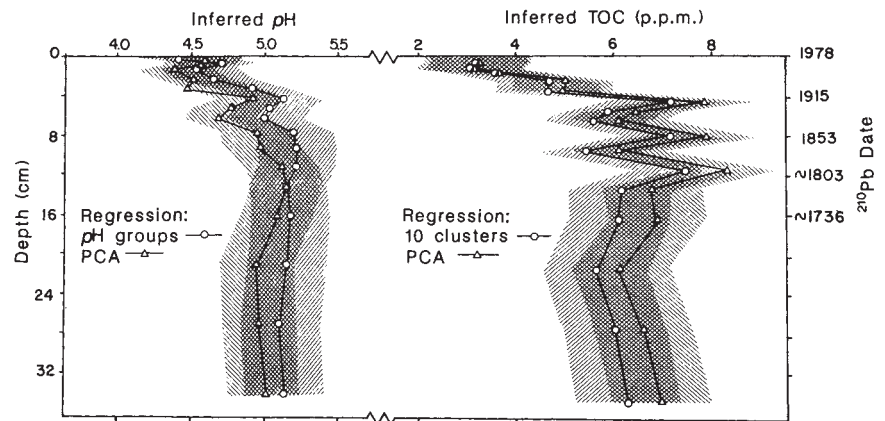
We applied the pH and TOC equations to two southern Norwegian lakes: (1) Hovvatn with a mean pH 4.4 (4.23–4.59, $n = 5$) and mean TOC 3.5 (2.7–3.9, $n = 3$) p.p.m.; and (2) Holmvatn with a mean pH 4.7 (4.53–4.74, $n = 5$) and mean TOC 2.2 (1.7–2.8, $n = 3$) p.p.m. based on sampling in 1974–79 according to refs 13 and 14 where additional water chemistry and physical data and locations are given. The fish populations disappeared from Hovvatn in the late 1940s and from Holmvatn in the 1960s⁸. Undisturbed sediment cores were obtained in 1978 (ref. 7). The ²¹⁰Pb constant-rate-of-supply model¹⁷ was used for dating sediment back to around 1880, beyond which dating was based on extrapolation parallel to the date–depth slope derived from the constant-initial-concentration model¹⁸.

The pH inferences for Hovvatn (Fig. 1) indicate that before the mid-1800s the lake was pH 4.9–5.2, having its least acidic period since roughly 1700. From the mid-1800s to around 1920 the inferred pH fluctuated between 4.7 and 5.1 and may have been slightly lower on average than before the mid-1800s. At around 1920 a more marked acidification began, resulting in inferred pHs of 4.4–4.6 by the 1950s. In the reconstruction based on principal components analysis (PCA), the acidification was completed in the 1920s; in the reconstruction based on pH groups it was completed in the 1940s. The total decreases were 0.4–0.8 of a pH unit, depending on the mode of calculation. The acidification inferences are due to the up-core changes in percentages of diatom taxa that strongly indicate decreasing pH⁴, including increases in the acidobiontic¹⁸ *Semiorbis hemicyclus* (Ehr.) Patr., *Tabellaria binalis* (Ehr.) Grun., and *Anomoeoneis serians* (Bréb. ex Kütz.) Cl., the acidophilic *Eunotia bactriana* Ehr., *Navicula subtilissima* Cl. and *N. heimantii* Van Dam et Kooyman, and decreases in the acidophilic *Cymbella perpusilla* A. Cl. and certain other *Navicula* species (such as, *N. mediocris* Krasske) (Fig. 2).

The TOC inferences for Hovvatn (Fig. 1) indicate that the lake was quite humic (TOC 6.1–8.3 p.p.m. (PCA) or 5.4–

* Present address: 2920 Leira, Valdres, Norway.

Fig. 1 pH and TOC inferences compared with depth and dates of sediment from Hovvatn, based on application of regression equations (ref. 3 and Table 1) to sedimentary diatom counts. Shading indicates the regression standard errors (S_e).



7.5 p.p.m. (clusters)) before about 1920, when a decrease started and continued until the late 1960s, reaching inferred values of 3.1–3.2 p.p.m. The total decreases relative to pre-acidification averages were 54% (PCA) and 48% (clusters). The decreases were inferred from changes in percentages of taxa whose present distributions are essentially unrelated to pH in the pH range 4.4–6.7 but are strongly related to TOC concentration⁴. These include up-core decreases in *Frustulia rhomboides* (Ehr.) DeT. and *Anomoeoneis seriens* v. *brachysira* (Bréb. ex Kütz.) Hust. (Fig. 2). In 1982, 1 yr after liming Hovvatn, a 68% increase was observed in the TOC that was not due to an increase in plankton²⁰.

We calculated *t*-statistics to compare the averages of the post-acidification values with pre-acidification values separately for each of the two reconstructions of TOC and each of the two reconstructions of pH (four cases). The average values of the top three core levels (TOC) or top four levels (pH) were compared with the average values of the bottom eight core levels for each parameter (Fig. 1). The null hypothesis was rejected ($P = 0.001$) in the four cases, indicating that significant changes took place in the inferred values of both parameters.

The reconstructions for Holmvatn (Fig. 3) reveal that before the 1940s the pH ranged from 4.9–5.1. In the 1940s, it started to decrease below this range, and by the 1960s had reached inferred pHs of 4.4–4.6 which continued to 1978. The total decrease in pH was 0.4–0.5 units. Diatom shifts responsible for these inferences are similar to those for Hovvatn (Fig. 2) (see ref. 7) except that corrections in taxonomic determinations for Holmvatn indicate that *Tabellaria quadrisepitata* increased from 0.3–2.0% before about 1960 to 4.5 and 7.5% in the top two levels of the core. A positive response of this taxon has been observed at acidified lochs in Scotland^{16,21}. The effect of the taxon as an acidobiont was not incorporated in the inferred pH profiles (Fig. 3) because it was not identified in the 35-lake data set on which the regression equations were based³. Had the taxon been included in pH inferences for Holmvatn, it would

have lowered the inferred pH by an estimated additional 0.1–0.2 unit since around 1960 and would have had little or no effect before that date. The taxon was not present in significant numbers at Hovvatn.

Before 1930, inferred TOC concentrations at Holmvatn averaged 8.4 (PCA) or 7.8 p.p.m. (clusters) (Fig. 3). At around 1930 TOC decreased to 4.4–5.6 (PCA) or 4.2–5.2 p.p.m. (clusters) and at about 1950 decreased again, resulting in an inferred TOC of ~3 p.p.m. which was maintained to 1978. The total decreases in TOC were 60–65%. These inferences are based on decreases in *Frustulia rhomboides* and *Anomoeoneis seriens* v. *brachysira*⁷. Decreases in inferred TOC began ~10 yr before the inferred decreases in pH, perhaps due to buffering by organics and Al during their 'titration' by strong acid. At Hovvatn, the decreases in these parameters were ostensibly contemporaneous, but the slow sediment accumulation rate at Hovvatn may have precluded detection of non-contemporaneity.

We ran *t*-tests for Holmvatn, comparing post- and pre-acidification averages separately for each of two reconstructions of TOC and each of two of pH (four cases). The average values of the top three core levels (TOC) or top two levels (pH) were compared with the bottom four and five levels, respectively. The results indicate that significant ($P = 0.001$) changes took place for both cases of each parameter (Fig. 3).

Our results support the hypothesis that concentrations of organic matter in the water decreased along with acidification as measured by pH. Assuming that most of the TOC was present as dissolved organic substances typical of coloured, natural waters, one can calculate, on the basis of the reported acid properties and dissociation behaviours of such organic substances^{22,23}, that the pre-acidification pHs of ~5.0 could have resulted from them alone. Increasing inputs of mineral acids from the atmosphere directly to the lake, and indirectly through the watershed, would have (1) protonated organic molecules, increasing their tendency to aggregate and precipitate^{24,25}, and (2) mobilized Al (ref. 26) and other metals (such as Fe²⁷) in

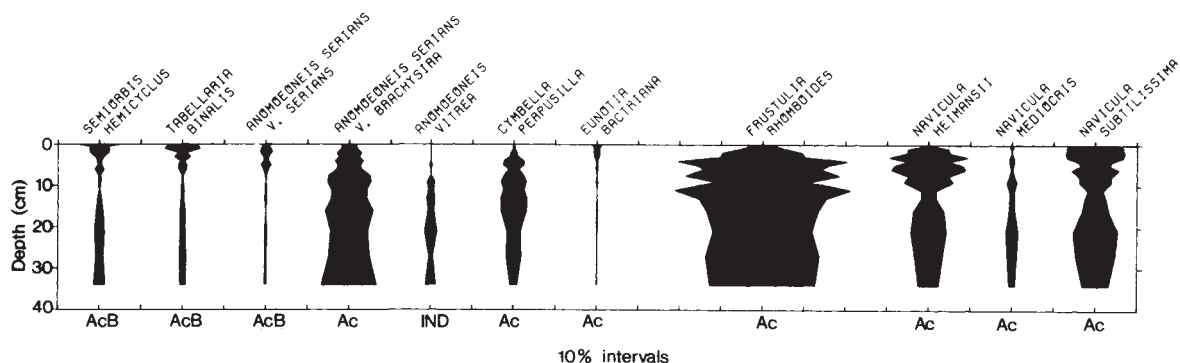


Fig. 2 Stratigraphy of important taxa in the Hovvatn core. AcB, acidobiontic; Ac, acidophilic; IND, indifferent (circumneutral)¹⁹.

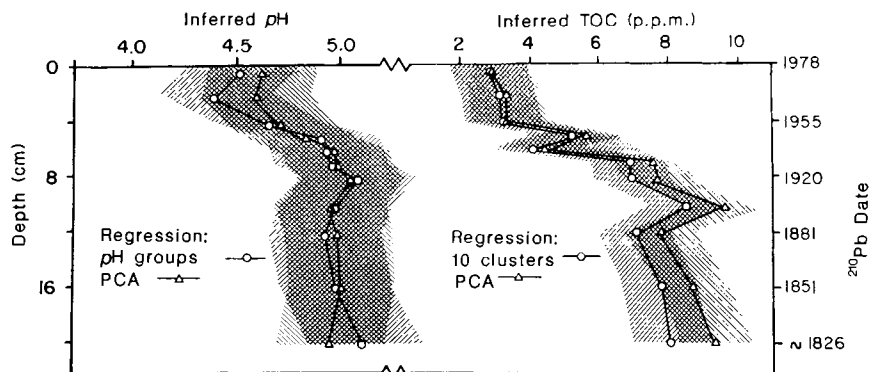


Fig. 3 pH and TOC inferences compared with depth and date of sediment from Holmvatn, based on application of regression equations (ref. 3 and Table 1) to sedimentary diatom counts. Shading indicates the regression standard errors (S_e).

inorganic form, providing further charge-neutralization of organic functional groups leading to their precipitation^{11,12,27}. Acidification of lakes such as Hovvatn and Holmvatn was not analogous to a titration of bicarbonate alkalinity, because such alkalinity was absent. Rather, it was a titration of organic matter, perhaps largely by acid-mobilized metals. Although the pH decreases were less than one unit in both lakes (but see 'Note added in proof' below), the resultant decrease in organics, along with increasing concentrations of strong mineral acids, effected a dramatic shift in water chemistry that was sufficient to eliminate the fishes. Before this shift, Al and trace metals would have been largely complexed with organics²⁸. After this shift, the proportion and total concentration of Al and perhaps other potentially toxic metals in free, inorganic form would have increased greatly²⁹. At both Hovvatn and Holmvatn, fishes finally succumbed⁸ 20–30 yr after rapid decreases in TOC began. Given the granitic lithology and extremely shallow, coarse soils topped by organic, mor layers in acidified lake districts in southern Norway¹³, naturally acidic (pH ~5), humic waters may have been the usual 'pre-acidification' condition in lakes there that subsequently lost their fish populations. Data from three acidified lakes in the Adirondack Mountains, New York, show changes in diatom stratigraphy (such as, decreases in *Frustulia rhomboides* and/or *Anomoeoneis seriata* v. *brachysira* coincident with increases in acidobiontic taxa), indicating that the changes in water chemistry we have postulated for Norway may also have occurred in the Adirondacks (refs 30, 31 and D. Charles, personal communication). This is supported by historical evidence for increased water transparency in Adirondack lakes¹.

Further palaeolimnological investigation of this acidification sequence is underway in North America, including the development of large data sets relating sedimentary remains of diatoms

to DOC and monomeric, inorganic Al (ref. 32). Direct investigations of Al and organic phases in the sediment could also provide useful information bearing on the acidification sequence. Additional insight would be gained by experimental research on responses of the key diatom taxa to changing pH, DOC, metals and other factors affected by acidification. While pH and Al have key roles in mechanisms of fish mortality³³, the causal connections between lake acidification and the observed shifts in diatom assemblages are unknown. In addition to investigations of the more or less direct chemical effects on diatoms of decreased pH, DOC, and metal complexation, work is needed on the indirect effects of the deepened euphotic zone, increased deep-water temperature, and altered depths and duration of thermal stratification that would result from decreased water colour and increased transparency^{11,12}.

We acknowledge the support of NSF grant DEB 78-10641 to R.B.D. and S. A. Norton, and US Department of Agriculture Hatch funding (ME 08465) to R.B.D. Personnel of the Norwegian Institute for Water Research, especially R. F. Wright, greatly facilitated this research. We thank M. L. Hoskins, S. T. Jagels and D. R. Beeson for their contributions to the diatom counts. For reviews and suggestions, we thank C. S. Cronan, M. (J.) Ford, G. L. Jacobson Jr, M. C. Whiting and C. T. Driscoll. S. A. Norton and his staff helped in several ways, and together with C. T. Hess and K. D. Hoagland provided the ²¹⁰Pb dates.

Note added in proof: The analysed sediment samples each represent a few years of deposit, and pH inferences based on them probably reflect some form of average pH for the period. Seasonal pH excursions such as spring flushes of low pH snow-melt are unresolved. Such 'shock depressions' in pH may have been a major cause of fish mortality³⁵, despite only small historical decreases in average (annual or longer) pH.

Received 21 January; accepted 5 June 1985.

- Schofield, C. *Trans. N.E. (U.S.) Fish Wildlife Conf.* 98–112 (1972).
- Almer, B., Dickson, W., Ekström, C. & Hörnström, E. *Ambio* 3, 30–36 (1974).
- Davis, R. B. & Anderson, D. S. *Hydrobiologia* 120, 69–87 (1985).
- Anderson, D. S., Davis, R. B. & Berge, F. *Hydrobiologia* (in the press).
- Battarbee, R. W. *Phil. Trans. R. Soc. B305*, 451–477 (1984).
- Charles, D. F. *Ecology* 66, 994–1011 (1985).
- Davis, R. B., Norton, S. A., Hess, C. T. & Brakke, D. F. *Hydrobiologia* 103, 113–123 (1983).
- Johannessen, M. & Wright, R. F. *SNFS Proj. Tekn. Not. (Oslo) TN53* (1980).
- Almer, B., Dickson, W., Ekström, C. & Hörnström, E. in *Sulfur in the Environment* (ed. Nriagu, J. O.) 271–311 (Wiley, New York, 1978).
- Dickson, W. in *Proc. int. Conf. Ecol. Impact Acid Precip.* (eds Drablos, D. & Tollan, A.) 75–83 (SNFS Project, Oslo, 1978).
- Effler, S. W., Schafran, G. C. & Driscoll, C. T. *Can. J. Fish. aquat. Sci.* (in the press).
- Yan, N. D. *Can. J. Fish. aquat. Sci.* 40, 621–626 (1983).
- Wright, R. F. *et al. SNFS Proj. Int. Rapp. (Oslo) 1R33* (1977).
- Henniksen, A. *SNFS Proj. Tekn. Not. (Oslo) TN46* (1979).
- Webb, T., III & Clark, D. R. *Annals N.Y. Acad. Sci.* 288, 93–118 (1977).
- Flower, R. J. & Battarbee, R. W. *Nature* 305, 130–133 (1983).
- Appleby, P. G. & Oldfield, F. *Catena* 5, 1–8 (1978).
- Robbins, J. A. in *The Biogeochemistry of Lead in the Environment* (ed. Nriagu, J. O.) 285–393 (Elsevier, Amsterdam, 1978).

- Hustedt, F. *Arch. Hydrobiol. Suppl.* 16, 274–394 (1939).
- Wright, R. F. *NIVA (Oslo) Acid Rain Res. Rep.* 5, (1984).
- Battarbee, R. W., Flower, R. J., Stevenson, A. C. & Rippey, B. *Nature* 314, 350–352 (1985).
- Oliver, B. G., Thurman, E. M. & Malcolm, M. L. *Geochim. cosmochim. Acta* 47, 2031–2035 (1983).
- Thurman, E. M. & Malcolm, R. L. in *Aqueous and Terrestrial Humic Materials* (eds Cristman, R. F. & Gjessing, E. T.) Ch. 1 (Ann Arbor, Michigan, 1983).
- Choppin, G. R. & Kullberg, L. *J. inorg. nucl. Chem.* 40, 651–654 (1978).
- Hayes, M. H. B. & Swift, R. S. in *The Chemistry of Soil Constituents* (eds Greenland, D. J. & Hayes, M. H. B.) 179–320 (Wiley, New York, 1978).
- Cronan, C. S. & Schofield, C. L. *Science* 204, 304–306 (1979).
- Schnitzer, M. & Khan, S. U. *Humic Substances in the Environment* (Dekker, New York, 1972).
- Driscoll, C. T., Jr, Baker, J. P., Bisogni, J. J. & Schofield, C. L. *Nature* 284, 161–164 (1980).
- Driscoll, C. T., Baker, J. P., Bisogni, J. J. & Schofield, C. L. in *Geological Aspects of Acid Deposition* (ed. Bricker, O. P.) Ch. 4 (Butterworth, Boston, 1984).
- Charles, D. F. *Verh. Int. Verein. Limnol.* 22, 559–566 (1984).
- Davis, R. B., Anderson, D. S., Charles, D. F. & Galloway, J. N. *Completion Report* (Electric Power Research Institute, Palo Alto, 1984).
- Electric Power Research Institute (EPRI), *PIRLA Proj.* (EPRI, Palo Alto, 1983).
- Dillon, P. J., Yan, N. D. & Harvey, H. H. *CRS Crit. Rev. Envir. Contr.* 13, 167–194 (1984).
- Dixon, W. J. & Brown, W. B. (eds) *Biomedical Computer Programs P-series* (University of California Press, Berkeley, 1979).
- Baker, J. P. in *The Acidic Deposition Phenomenon and its Effects* Vol. 2 (eds Altschuler, A. P. & Linthurst, R. A.) Ch. 5.6 (USEPA-600/8-83-016BF, Washington, DC, 1984).

Category-specific naming deficit following cerebral infarction

John Hart Jr*, Rita Sloan Berndt†‡
& Alfonso Caramazza*

* Department of Psychology, The Johns Hopkins University, Baltimore, Maryland 21218, USA

† Department of Neurology, University of Maryland Hospital, 22 S. Greene Street, Baltimore, Maryland 21201, USA

Studies aimed at characterizing the operation of cognitive functions in normal individuals have examined data from patients with focal cerebral insult. These studies assume that brain damage impairs functions of the cognitive processes along lines that honour the 'normal' pre-morbid organization of the cognitive system¹. For example, detailed study of individual brain-damaged patients has revealed apparently selective disruption of cognitive functions such as auditory/verbal working memory², phonological processing ability³, grapheme-to-phoneme translation procedures⁴ and semantic processing⁵. Warrington *et al.* have studied patients with even more fine-grained selective disturbances of the semantic system^{6,7}. The most selective deficits have been reported for four patients who were significantly better at identifying inanimate objects than they were at identifying living things and foods⁸. These patterns of selective deficit after localized brain damage provide important information about the normal organization of the lexicon, and ultimately about how components of the lexical system are related to particular neural substrates. Here, we report a case study of a patient demonstrating a very selective disturbance of the ability to name items from two related semantic categories. Despite normal performance on a large battery of lexical/semantic tasks, the patient shows a consistent and striking disability in naming members of the semantic categories of 'fruits' and 'vegetables'. The selectivity of this deficit supports a category-specific organization of the mental lexicon, and suggests independence of the processing routes involving naming and name recognition.

The patient studied here (M.D.) is a 34-year-old, right-handed male college graduate who works as a systems analyst for a large United States government agency. In August of 1981 he suffered a left-hemisphere cerebrovascular accident, which resulted acutely in a global aphasia and right hemiparesis. Within 1 month, he recovered to a mild expressive aphasia and mild hemiparesis. M.D. subsequently experienced several transient ischaemic attacks, and a left internal carotid artery occlusion was diagnosed. In December 1981, he was treated with extra-cranial/intra-cranial bypass surgery, which has been successful in averting further ischaemic episodes. A computerized tomography scan obtained at 1 month post-onset revealed an infarction involving the left frontal lobe and basal ganglia.

Early in 1983 M.D. was referred as a potential candidate for a research project on naming deficits in aphasia. His initial performance on all subtests of a standard language battery⁹ indicated that he was not clinically aphasic, although he had some difficulty naming objects. He maintained that he was experiencing considerable difficulty with certain words. Further experimental testing suggested that these difficulties were focused on the semantic categories of 'fruits' and 'vegetables'. M.D. showed a striking inability to name such common items as Peach and Orange while able to name easily less frequent items such as Abacus and Sphinx.

To chart the boundaries and stability of this unusual deficit, M.D. was tested over 1 year with a variety of materials that were designed to evaluate the structure of his lexical/semantic system. Performance on a large battery of tests was excellent. This battery included: visual and auditory lexical decision; oral

reading; word-picture matching (nouns and verbs); semantic categorization (words and pictures); and picture naming (oral and written). Unless some member of the semantic categories of fruits and vegetables were a stimulus item, M.D.'s scores were almost perfect.

Special tests were designed to focus on the problematic categories. M.D. was asked to name a large number of items (line drawings, coloured drawings, colour photographs and actual objects) from several semantic classes; Table 1 summarizes his performance for all stimulus types.

M.D. showed considerable difficulty in naming individual fruits and vegetables, but was able to name easily a large range of other pictures and objects. His seven naming errors outside the fruits/vegetables category involved two household items, four geometric shapes (for example, diamond) and one tree. He named correctly a total of 13 food products outside the categories of fruits and vegetables. The 'other' items given to M.D. for naming were chosen to cover a wide range of semantic categories. Some of these categories (for example, 'body parts') have members with a very high frequency of usage in the language, hence the mean frequency of occurrence¹⁰ for the fruit/vegetable items (11.8 per million) was considerably lower than the mean for the other items (32.8 per million). This disparity does not explain M.D.'s selective impairment, however, as 73% of items in the other group were in the same frequency range as the fruit/vegetable items and were named without difficulty.

To investigate M.D.'s knowledge of the semantic category of these items, he was shown pictures of 75 items from the categories fruits, vegetables, animals, vehicles and food products and he was asked to sort them into piles on the basis of their semantic classification. He was then asked to label the categories thus produced; his errors consisted entirely of confusions involving the categories of fruits and vegetables. He categorized 3 out of 24 fruits as vegetables and 6 out of 23 vegetables as fruit. He also incorrectly classified two food products (butter, cheese) as vegetables, although he named them correctly. Although the absolute number of errors committed on this task was not very large, M.D. had considerable difficulty with it, performing slowly and complaining of uncertainty.

In another attempt to determine M.D.'s knowledge of the members of semantic categories, he was asked to generate as many names as possible from 17 categories. The mean number of fruits and vegetables generated in 1 min was 6.5, with an additional four vegetable names generated incorrectly as fruits. The mean number of items generated for the other 15 categories was 12.4, including 12 instances in the category food products.

M.D.'s ability to name and to categorize pictures of fruits and vegetables is compromised relative to his ability to name and to categorize members of other categories. A further set of tasks was designed to assess the possibility that this impairment is limited to stimuli processed through the visual modality. A set of 20 verbal definitions, containing perceptual, functional and category information, was developed for 10 fruit/vegetable items and 10 other items (animals, furniture and clothing). M.D. named two out of 10 fruits and vegetables and all of 10 other items from their definitions. Next, M.D. was asked to name a

Table 1 Number of correct naming responses

	Semantic category		
	Fruit	Vegetables	Other
Line drawings	5/11	7/11	11/11
Coloured drawings	4/6	5/7	18/18
Photographs	11/18	12/18	222/229
Real objects	10/13	13/23	11/11
Total	30/48 (0.63)	37/59 (0.63)	262/269 (0.97)

The 'other' category includes vehicles, toys, tools, animals, body parts, food products, school, bathroom, kitchen and personal items, clothing, colours, shapes and trees.

‡ To whom correspondence should be addressed.

set of objects that he could feel with either the left or right hand, but could not see. Fruit/vegetable items were selected on the basis of their tactile discriminability from other similar items. M.D. named 6 of 13 fruits, 11 of 21 vegetables and 11 of 12 other items (he failed on tea bag). There was no difference in performance based on the hand used.

M.D.'s selective naming deficit for fruit/vegetable items is not modality-specific, and therefore may be the result of a selective disturbance within the semantic system itself, rather than a problem with access to the semantic system. To assess this possibility, a series of tasks was designed to probe M.D.'s comprehension of the names that he has difficulty producing. A 45-item word/picture matching task was developed in which M.D. was required to point to one of two pictures in response to an aurally presented word. Fruit and vegetable items were represented, as well as vehicles, food products and animals. For half the items, the incorrect pictures were closely semantically related to the target. M.D.'s one error (confusion between rhinoceros and hippopotamus) was not within the fruits/vegetable category; he pointed immediately and with certainty to pictured fruits and vegetables on hearing their names.

M.D.'s comprehension of the properties of objects was assessed by asking him for judgements about the category, size, colour, texture and shape of eight fruit/vegetable items that he had previously misnamed, and four animals. Item names were presented aloud. Although some responses on this task were hesitant for the fruits/vegetables categories, they were correct for all properties.

Note that M.D. could categorize items correctly as fruits or vegetables when their names were presented aurally, suggesting that his difficulty in categorizing pictures of fruits and vegetables was related to his inability to name them. This possibility was supported by his ability to categorize correctly all the written names of the fruits and vegetables whose pictures he had found difficult to classify. On two separate occasions approximately 6 months apart, M.D. easily sorted 14 printed fruit/vegetable names into their proper categories, together with the names of 21 items from other categories.

The impairment to the semantic system revealed in this patient appears to be limited to two specific and related semantic categories and to situations that require him to name the objects or to categorize them without having first been given their names. M.D. is aware of this problem, and expresses some frustration with it. Although he confesses to knowing little about cooking or food in general, it is clear that his difficulty with these items is not consonant with his ability to name many other types of foods and food-related items; in this respect he differs from the patients studied by Warrington and Shallice⁸, who had difficulty identifying all types of food products. Further, his failure to learn the names of these items after many test sessions that focused on this deficit suggests that this phenomenon does not reflect a simple pre-morbid lack of interest in these categories.

Although a unified lexical/semantic theory cannot be formulated on the basis of this one case, there are three important implications of these findings for the ultimate development of such a theory. First, the selective impairment of information in specific superordinate categories suggests that the organization of the semantic system in some sense honours those categorical distinctions. These results support and considerably extend previous neuropsychological investigations which have indicated a category-specific organization of the semantic system^{6,8}. Second, the dissociation in categorization ability between performance with lexical instances (which is normal) and with pictorial instances (which is impaired) suggests that lexical categorization could be accomplished on the basis of strictly lexical, as opposed to semantic, information. Third, although a general dissociation between 'name recognition' and 'name retrieval' has been supported previously by results from aphasic patients⁹, the category-specific dissociation found in M.D. indicates that the output lexicon is addressed by semantically categorized information that can be disrupted highly selectively.

The results reported here suggest that the lexical/semantic

system is organized categorically and specifically at the level of the input and output processes to and from the system.

This study was supported by NINCDS grants NS-21054 to the University of Maryland School of Medicine and NS-14099 to The Johns Hopkins University. We thank Charlotte Mitchum for assistance with many aspects of this research.

Received 3 December 1984; accepted 2 April 1985.

1. Saffran, E. M. *Br. J. Psychol.* **73**, 317 (1982).
2. Warrington, E. K. & Shallice, T. *Brain* **92**, 885 (1969).
3. Caramazza, A., Berndt, R. S. & Basili, A. *Brain Language* **18**, 128 (1983).
4. Beauvois, M. F. & Derousne, J. J. *Neurol. Neurosurg. Psychiat.* **42**, 1115 (1979).
5. Schwartz, M. F., Marin, O. S. M. & Saffran, E. M. *Brain Language* **7**, 277 (1979).
6. Warrington, E. *Br. J. Psychol.* **72**, 175-196 (1981).
7. Warrington, E. & McCarthy, R. *Brain* **106**, 859 (1983).
8. Warrington, E. & Shallice, T. *Brain* **107**, 829 (1984).
9. Goodglass, H. & Kaplan, E. *The Assessment of Aphasia and Related Disorders* (Lea & Febiger, Philadelphia, 1972).
10. Francis, N. & Kucera, H. *Frequency Analysis of English Usage* (Houghton Mifflin, Boston, 1962).

Three types of neuronal calcium channel with different calcium agonist sensitivity

Martha C. Nowicky*‡, Aaron P. Fox†§
& Richard W. Tsien†

* Section of Neuroanatomy and † Department of Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

How many types of calcium channels exist in neurones? This question is fundamental to understanding how calcium entry contributes to diverse neuronal functions such as transmitter release, neurite extension, spike initiation and rhythmic firing¹⁻³. There is considerable evidence for the presence of more than one type of Ca conductance in neurones⁴⁻¹³ and other cells¹⁴⁻¹⁸. However, little is known about single-channel properties of diverse neuronal Ca channels, or their responsiveness to dihydropyridines, compounds widely used as labels in Ca channel purification¹⁹⁻²¹. Here we report evidence for the coexistence of three types of Ca channel in sensory neurones of the chick dorsal root ganglion. In addition to a large conductance channel that contributes long-lasting current at strong depolarizations (L), and a relatively tiny conductance that underlies a transient current activated at weak depolarizations (T), we find a third type of unitary activity (N) that is neither T nor L. N-type Ca channels require strongly negative potentials for complete removal of inactivation (unlike L) and strong depolarizations for activation (unlike T). The dihydropyridine Ca agonist Bay K 8644 strongly increases the opening probability of L-, but not T- or N-type channels.

Figure 1 shows evidence for three distinct components of Ca channel current in whole-cell recordings obtained under ionic conditions that minimize contamination by other currents. The components were distinguished kinetically by applying depolarizing test pulses at various levels from different holding potentials (h.p.). With h.p. = -40 mV (uppermost traces in each panel), strong depolarizations (-10, +10, +20 mV) are required to activate any inward current; the peak current-voltage relation (Fig. 1b, squares) is typical for a single-current component. Because this inward current component decays very slowly ($t_{1/2}$ of hundreds of milliseconds), we designate it 'L' (for long-lasting). Weak depolarizations from h.p. = -100 mV evoke a different Ca channel current (Fig. 1a), seen as a prominent shoulder at negative test potentials in the associated peak current-voltage plot (Fig. 1b, circles). The additional current becomes noticeable at -60 mV, and is nearly constant in amplitude between -40 and -10 mV, consistent with a very negative range of activation. As this component decays relatively rapidly

‡ Present address: Department of Anatomy, The Medical College of Pennsylvania, Philadelphia, Pennsylvania 19129, USA

§ To whom reprint requests should be addressed.

Fig. 1 Three components of Ca channel current in whole-cell recordings with 10 mM extracellular Ca. *a*, Superimposed current traces evoked from h.p. = -40 and h.p. = -100 mV. Test potential indicated above. Linear leak and capacitance have been subtracted with 1 ms blanked at both the on and off of the depolarizing pulse. *b*, Plot of peak current versus test potential from h.p. = -40 (squares) and h.p. = -100 mV (circles). *c*, Plot of magnitude of relaxing current against test potential from h.p. = -100 mV. The current relaxation is the late current subtracted from the peak current. Inset repeats records obtained at +10 mV and shows how peak and relaxing current amplitudes were measured. Cell N82F. **Methods.** Dorsal root ganglia were obtained from 8-12-day-old chick embryos. The isolation and culture procedures followed standard protocols (see ref. 24). Cells were maintained in culture for 2-7 days before recording. The internal solution (pipette) contained (in mM): 100 CsCl, 10 Cs-EGTA, 5 MgCl₂, 40 HEPES, 2 ATP, 0.25 cyclic AMP, pH 7.3. The external solution was exchanged after formation of a gigaseal from a standard Tyrode solution to one containing (in mM) 10 CaCl₂, 135 TEA-Cl, 10 HEPES, and 200 μ M tetrodotoxin (TTX), pH 7.3. Voltage pulses and current traces were simultaneously generated and sampled online with a PDP 11/23 computer. Whole-cell currents were filtered with an 8-pole Bessel filter (3 dB down at 3 kHz) and sampled at 190- μ s intervals. The stimulation rate for both whole-cell and single-channel experiments was 0.25 Hz. All experiments were carried out at room temperature (21-22 °C).

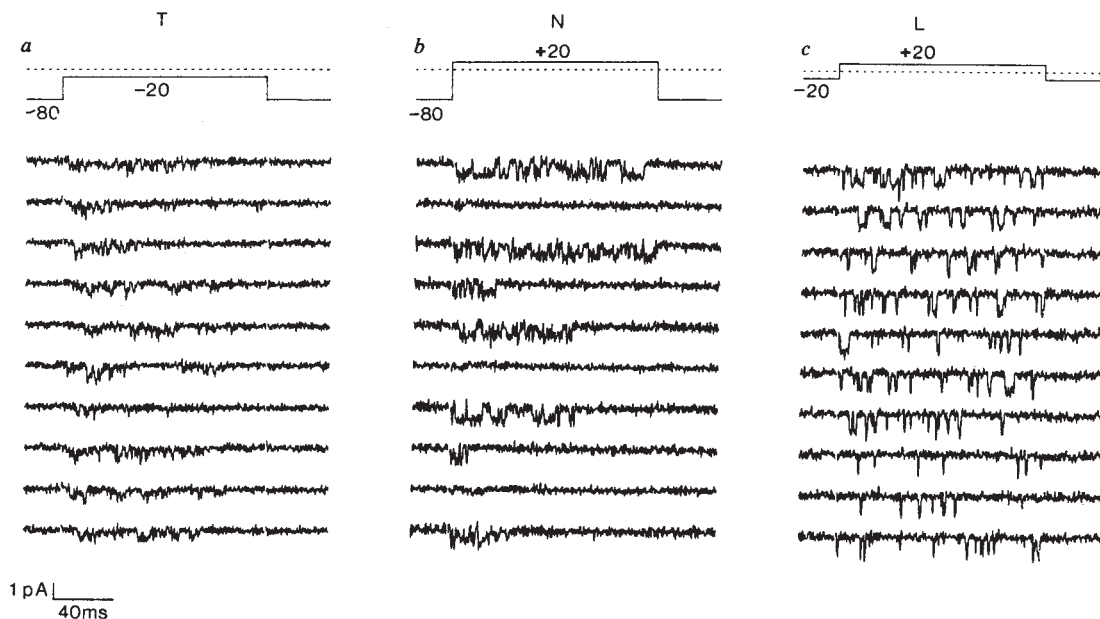
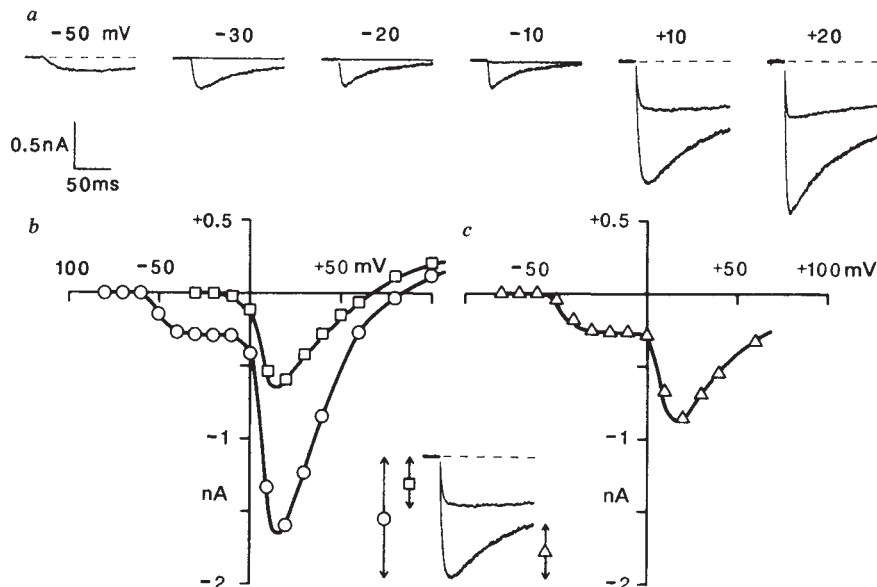


Fig. 2 Three types of unitary Ca channel activity seen in cell-attached patch recordings with Ba as the charge carrier. *a*, T-type channel activity, cell N85A. *b*, N-type channel, cell N75H. *c*, L-type channel, cell N67J. Patch pipettes contained (in mM): 110 BaCl₂, 10 HEPES, and 200 nM TTX, pH 7.4. To put patch membrane potential on an absolute scale, cell resting potential was zeroed with an external solution containing (mM): 140 K-aspartate, 10 K-EGTA, 10 HEPES, 1 MgCl₂, pH 7.4. Current signals were filtered at 1 kHz and sampled at 5 kHz. Within each panel traces are consecutive. Values of single-channel amplitude and slope conductances given in the text were obtained from amplitude histograms. The unitary current amplitude of the T-type channel at -20 mV was given by the average of amplitudes from three different patches exhibiting long, well-resolved openings. We have seen each of the three channel types in isolation and in all possible combinations. The number of channels under the patch pipette varied considerably from one to many.

($\tau \sim 25$ ms at -30 mV), we term it 'T' (transient). A third component of calcium current appears with strong depolarizations from h.p. = -100 mV (for example, the traces at +10 or +20 mV in Fig. 1*a*). It is distinguished most easily in a plot of the amplitude of decaying current versus test potential (triangles, Fig. 1*c*). The decay amplitude shows a clear plateau between -30 and -10 mV, as expected for component T; however, it grows substantially with stronger test pulses, reaching a peak at +20 mV. We attribute the extra decaying current to a third component that we term 'N' (neither T nor L). Like T, but unlike L, component N contributes phasic current and requires strongly

negative holding potentials for complete removal of inactivation. Like L, but unlike T, N current requires strong depolarizations for activation and is relatively sensitive to the inorganic blocker cadmium. For example, in six cells, 50 μ M Cd almost completely abolished N and L current, but left $55 \pm 4\%$ of T (see ref. 22). All three components are present in most whole-cell recordings, although the relative amplitudes vary considerably between cells and with different ionic conditions. N current is particularly prominent with 55 or 110 mM Ba as the charge carrier; in whole-cell recordings from four cells (average capacitance 30 pF), the maximal amplitude of the decaying N-type current

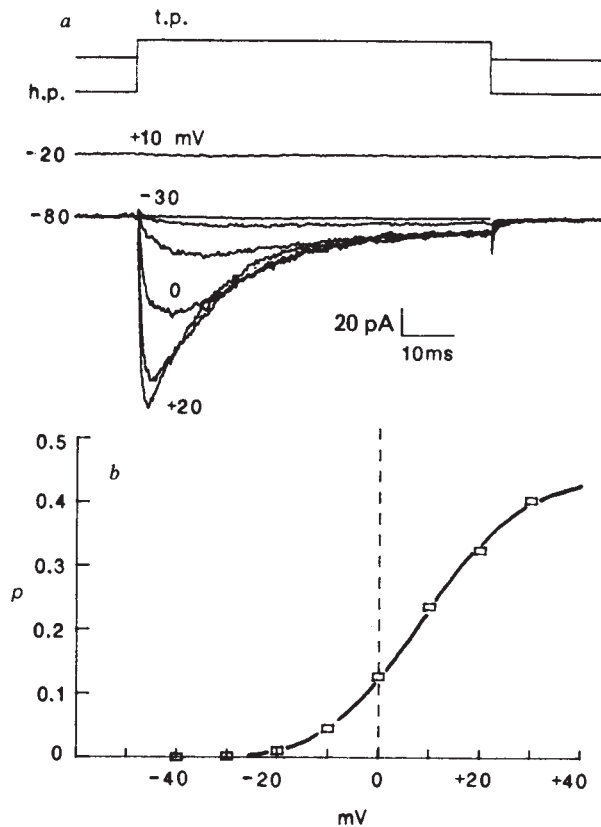


Fig. 3 Voltage-dependent activation of N-type Ca channel activity in a cell-attached patch with ~ 600 N-type channels. *a*, Voltage-clamp protocols (top); associated current records produced by depolarization from h.p. = -20 to test potential (t.p.) = $+10$ mV (middle trace) and by depolarizations from h.p. = -80 to test potentials ranging from -30 mV to $+20$ mV in 10 -mV increments (bottom family of traces). Each record was obtained by averaging five individual sweeps after subtraction of linear leak and capacity currents. *b*, Analysis of voltage-dependence of peak opening probability (p) from the same experiment. Values of p were obtained by dividing the peak current ($= N_T \times p \times i$) by the unitary current (i) at each test potential and by an estimate of the total number of channels ($N_T = 599$). Unitary current was given by the equation $i(V) = -1.023 \text{ pA} + (13.2 \text{ pS}) \times (V)$ single-channel data from 18 cell-attached patches (correlation coefficient for linear regression was greater than 0.99). N_T was determined by comparison with the single-channel experiment in Fig. 2b, which showed that $p = 0.32$ at $+20$ mV. Patch 84G.

at $+20$ mV was $1,340 \pm 236$ pA, roughly equal to the L current at $+20$ mV and ~ 8 times greater than the peak T current at -10 mV.

Recordings from cell-attached patches on dorsal root ganglion (DRG) cell bodies show three types of unitary Ca channel current, readily distinguished by different slope conductances in 110 mM Ba (Fig. 2). The different channel types show kinetic properties that correspond well to components T, N and L in whole-cell recordings. Figure 2a shows unitary current activity of a channel that we associate with component T. Inactivation of this channel is complete positive to -40 mV, and is fully removed only at holding potentials more negative than -80 mV. Activation becomes detectable positive to -50 mV. With test depolarizations to -20 or -10 mV, the averaged current decays with a time course very similar to component T in whole-cell recordings. The T-type channel has a small slope conductance (about 8 pS in 110 mM Ba).

Figure 2c illustrates unitary channel activity which corresponds to the L component. It is relatively resistant to inactivation: activity is seen at h.p. = -20 mV (Fig. 2c) or even at 0 mV. Activation of the L-type channel usually becomes significant at -10 mV, although openings can sometimes be detected as nega-

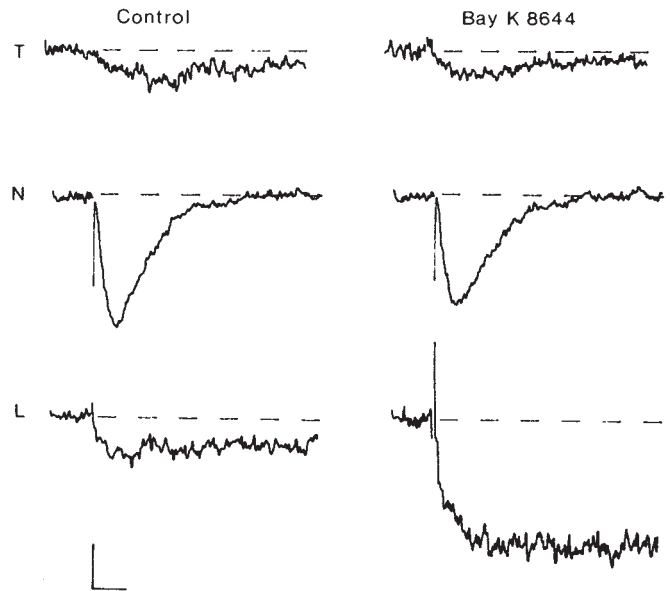


Fig. 4 Single-channel current averages recorded before and after the addition of Bay K 8644. All recording conditions identical to those described in Fig. 2 legend. Averaged currents for three channel types before (left) and after exposure to $5 \mu\text{M}$ Bay K 8644 (right). Top, T-type activity, voltage step from -80 to -30 mV. Patch 86D. Middle, N-type activity, voltage step from -80 to $+10$ mV. Patch 84K. Bottom, L-type activity, voltage step from -40 to $+10$ mV. Patch 89F. Vertical calibration bar represents 1 pA for middle panel and 0.25 for top and bottom panels. Horizontal calibration bar represents 20 ms for top and bottom panels and 10 ms for middle panel.

tive as -30 mV. Averaged currents obtained from many individual traces do not relax at any test potential (see Fig. 4). The channel is characterized by a large slope conductance in 110 Ba (25 pS).

Figure 2b illustrates a third type of unitary activity that we term 'N' because its properties account for component N in whole-cell recordings. N-type Ca channels differ from L- and T-type Ca channels in several respects. Their slope conductance with 110 mM Ba is 13 pS, larger than that of T-type channels and smaller than that of L-type channels. The difference in unitary current size is highly significant: for example, at -20 mV, average values ($\pm$ s.e.m.) for T, N and L were 0.62 ± 0.03 (5 patches), 1.22 ± 0.03 (10 patches) and 2.07 ± 0.09 (6 patches), respectively. At $+20$ mV, the test potential illustrated in Fig. 2b and c, the difference between average values for N and L (0.824 and 1.12 pA) is smaller in absolute terms but is still appreciable.

N-type Ca channels can also be distinguished from L-type Ca channels by their inactivation properties. The time-dependence of inactivation of N-type Ca channels appears as a bunching of openings near the beginning of the depolarization and an absence of openings near the end (Fig. 2b); the averaged current record (not shown) decays almost completely by the end of a 136 -ms test pulse. The steady voltage-dependence of inactivation was studied by varying the holding potential. Unitary N-type activity is completely abolished by holding the patch at -20 mV, a potential at which L-type channels remain largely available for opening (see Fig. 2c). N- and T-type Ca channels are similar in their inactivation properties but very different in the voltage-dependence of activation. With isotonic Ba in the pipette, significant activation of N-type channels requires depolarization to -20 mV, whereas opening of T-type channels first becomes detectable beyond -50 mV.

In several experiments, we recorded from cell-attached patches with many N-type Ca channels and no detectable T- or L-type activity. Figure 3 shows kinetic analysis from this kind of recording. The rise and fall of the N-type channel current both become faster as the test pulse depolarization is increased;

at strong depolarizations, inactivation is largely complete within ~100 ms (Fig. 3a, bottom family of traces). A steady holding potential of -20 mV also produces complete inactivation (Fig. 3a, middle trace); 50% inactivation is seen at h.p. = -60 mV (trace not shown). As Fig. 3b illustrates, the open probability (p) increases over the range between -20 and +30 mV, in good agreement with recordings from whole cells (Fig. 1c) and patches with only one or a few N-type channels (Fig. 2b).

In addition to differences in kinetics, unitary conductance and sensitivity to Cd, we find that the three channel types differ in their responsiveness to Bay K 8644. This dihydropyridine Ca agonist²³ elevates calcium influx by promoting a pattern of Ca channel gating with very long openings in neurones²⁴ and heart cells²⁵⁻²⁸. Bay K 8644 strongly enhances averaged L-type channel currents in DRG neurones^{11,24} (Fig. 4), but produces no change in the current carried by T-type channels (6 patches) or N-type channels (4 patches). Our results with Bay K 8644 are consistent with findings of dihydropyridine-resistant Ca channels in GH3 cells¹⁸ and cardiac cells²⁹. The existence of different dihydropyridine-resistant channels in neurones might help explain the partial or complete lack of effect of dihydropyridines on depolarization-induced Ca entry or transmitter release in many systems^{30,31}.

Although this is the first report to focus on the coexistence of three types of calcium channels in one preparation, we believe that activity of all three channel types may be found in whole-cell currents reported by others (for example Fig. 1c of ref. 8). Single-channel recordings in DRG neurones by Carbone and Lux⁹ have shown two types of unitary activity that correspond kinetically to what we call T-type and L-type Ca channels. The relative sizes of the unitary currents with 40 mM Ca as the charge carrier are opposite to what we find with 110 mM Ba; this discrepancy might be explained by the very different ionic selectivity properties of T- and L-type Ca channels^{10,17,29}.

It is attractive to suppose that the various channels may serve different cellular functions. For example, T-type channels might contribute to threshold behaviour or rhythmic activity and N- or L-type channels may be important for dendritic spikes or neurotransmitter release^{3-5,11,17}. Determination of the cellular distribution of the channel types and their sensitivity to neuromodulators will help in the elucidation of their physiological roles.

We thank So-Ching Wong for preparing the cultured neurones and for help with the illustrations, and E. W. McCleskey, P. Hess, B. Nilius and J. B. Lansman for helpful discussion. Support was provided by a Canadian Heart Foundation Fellowship to A.P.F. and USPHS grants to M.C.N. & R.W.T.

Received 7 February; accepted 22 May 1985.

1. Miller, R. J. *Trends Neurosci.* **8**, 45-47 (1985).
2. Hagiwara, S. & Byerly, L. *Trends Neurosci.* **6**, 189-193 (1983).
3. Llinas, R., Steinberg, I. Z. & Walton, K. *Biophys. J.* **33**, 323-352 (1981).
4. Fishman, M. C. & Spector, I. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5245-5249 (1981).
5. Llinas, R. & Sugimori, M. *J. Physiol., Lond.* **305**, 197-213 (1980).
6. Murase, K. & Randic, M. *J. Physiol., Lond.* **344**, 141-153 (1983).
7. Nowycky, M. C., Fox, A. P. & Tsien, R. W. *Biophys. J.* **45**, 36a (1984).
8. Carbone, E. & Lux, H. D. *Biophys. J.* **46**, 413-418 (1984).
9. Carbone, E. & Lux, H. D. *Nature* **310**, 501-502 (1984).
10. Yoshii, M., Tsunoo, A. & Narahashi, T. *Biophys. J.* **47**, 433a (1984).
11. Brown, D. A. et al. *IUPHAR Proc.* (in the press).
12. Fedulova, S. A., Kostyuk, P. G. & Veselovsky, N. S. *J. Physiol., Lond.* **359**, 431-446 (1985).
13. Bossu, J.-L., Feltz, A. & Thomann, J.-M. *Pflügers Arch. ges. Physiol.* (in the press).
14. Hagiwara, S., Ozawa, S. & Sand, O. *J. gen. Physiol.* **65**, 617-644 (1975).
15. Deitmer, J. W. *J. Physiol., Lond.* **355**, 137-159 (1984).
16. Fox, A. P. & Krasne, S. J. *J. Physiol., Lond.* **356**, 491-505 (1984).
17. Armstrong, C. M. & Matteson, D. R. *Science* **227**, 65-67 (1985).
18. Cohen, C. J. & McCarthy, R. T. *Biophys. J.* **47**, 513a (1985).
19. Curtis, B. M. & Catterall, W. A. *J. biol. Chem.* **258**, 9344-9348 (1983).
20. Norman, R. I., Borsotto, M., Fosset, M. & Lazdunski, M. *Biochem. biophys. Res. Commun.* **111**, 878-883 (1983).
21. Glossman, H., Ferry, D. R., Lubbecke, F., Mewes, R. & Hoffman, F. *Trends pharmac. Sci.* **3**, 431-437 (1982).
22. Nowycky, M. C., Fox, A. P. & Tsien, R. W. *Soc. Neurosci. Abstr.* **10**, 526 (1984).
23. Schramm, M., Thomas, G., Towart, R. & Franckowiak, G. *Nature* **303**, 535-537 (1983).
24. Nowycky, M. C., Fox, A. P. & Tsien, R. W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2178-2182 (1985).
25. Kokubun, S. & Reuter, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4824-4827 (1984).
26. Ochi, R., Hino, N. & Niimi, Y. *Proc. Jap. Acad.* **B60**, 153-156 (1984).
27. Hess, P., Lansman, J. B. & Tsien, R. W. *Nature* **311**, 538-544 (1984).
28. Brown, A. M., Kunze, D. L. & Yatani, A. *Nature* **311**, 570-572 (1984).
29. Nilius, B., Hess, P., Lansman, J. B. & Tsien, R. W. *Nature* **316**, 443-446 (1985).
30. Nachsen, D. A. & Blaustein, M. P. *Molec. Pharmac.* **16**, 579-586 (1979).
31. Turner, T. J. & Goldin, S. M. *J. Neurosci.* **5**, 841-849 (1985).

A novel type of cardiac calcium channel in ventricular cells

B. Nilius

Julius Bernstein Institute of Physiology, Martin Luther Universität, 4020 Halle (Saale), GDR

P. Hess, J. B. Lansman* & R. W. Tsien†

Department of Physiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA

Calcium influx is vital for several aspects of cardiac activity, so it is important to ask if heart cells possess a single or multiple types of Ca channel. Only one Ca channel type has been identified in patch-clamp studies of unitary current¹⁻³, despite suggestions to the contrary from whole-cell recordings in heart cells⁴⁻⁶ and unitary recordings from other cells^{7,8}. Here we describe a novel type of cardiac Ca channel with several properties that distinguish it from the hitherto-identified Ca channel in heart cells. Its conductance in isotonic Ba is small (8 pS), and is no larger in Ba than in Ca. It activates and inactivates at relatively negative potentials and remains functional long after patch excision. It is insensitive to dihydropyridines such as nimodipine and the Ca agonist Bay K 8644, and is more resistant to block by external Cd than the previously described type of cardiac Ca channel.

Figure 1 illustrates some of the features that distinguish the two types of Ca channel. Activity of the previously described type of cardiac Ca channel (Fig. 1a) was evoked by depolarizing voltage-clamp pulses to +10 mV from a holding potential of -50 mV; with 110 mM Ba in the pipette, channel openings produce current pulses of ~1.2 pA amplitude. The pattern of activity is typified by brief openings occurring in clusters throughout the pulse, with occasional sweeps dominated by long-lasting openings. The averaged current is fairly well-maintained. Figure 1b shows strikingly different channel activity, recorded from the same patch with pulses from a holding potential of -70 mV to -20 mV, a test potential that is normally too negative with 110 mM Ba in the pipette to activate the hitherto-identified channel. Here openings appear as inward current pulses of ~0.5 pA amplitude, and are generally grouped as one burst of activity that tends to occur near the beginning of each pulse; the averaged current peaks at ~10 ms and then decays rapidly. As the unitary current is larger in Fig. 1a than in b, despite the smaller driving force for Ba entry at the more positive test potential, there can be little doubt of the existence of two different kinds of channel. The slope conductance between -50 and 0 mV is 8 pS for the channel type in Fig. 1b, considerably smaller than the conductance of the type in a (18-25 pS in ~100 mM Ba, from our own studies and refs 1, 2). The two channels also differ in their ability to conduct Ca or Ba. Unlike the previously described Ca channel, which supports a much larger unitary flux with Ba than with Ca (compare Figs 1a and 3c, d), the second type of Ca channel shows very similar single-channel currents with either ion (Fig. 1b, c). The time course of the averaged currents is also very much the same.

As suggested by the voltage protocols in Fig. 1, we found major differences in the voltage-dependence of the two channel types. With 110 mM Ba in the pipette, the second type of Ca channel showed detectable activation and 90% complete steady-state inactivation near -50 mV, about 40 mV negative to a corresponding potential for the previously described type of Ca channel.

Although the voltage dependence of the new type of channel qualitatively resembles that of Na channels, the unitary events cannot be explained by Ca or Ba movements through Na channels: they persist in the presence of doses of tetrodotoxin (TTX,

* Present address: Physiological Laboratories, Downing Street, Cambridge CB2 3EG, UK.

† To whom reprint requests should be addressed.

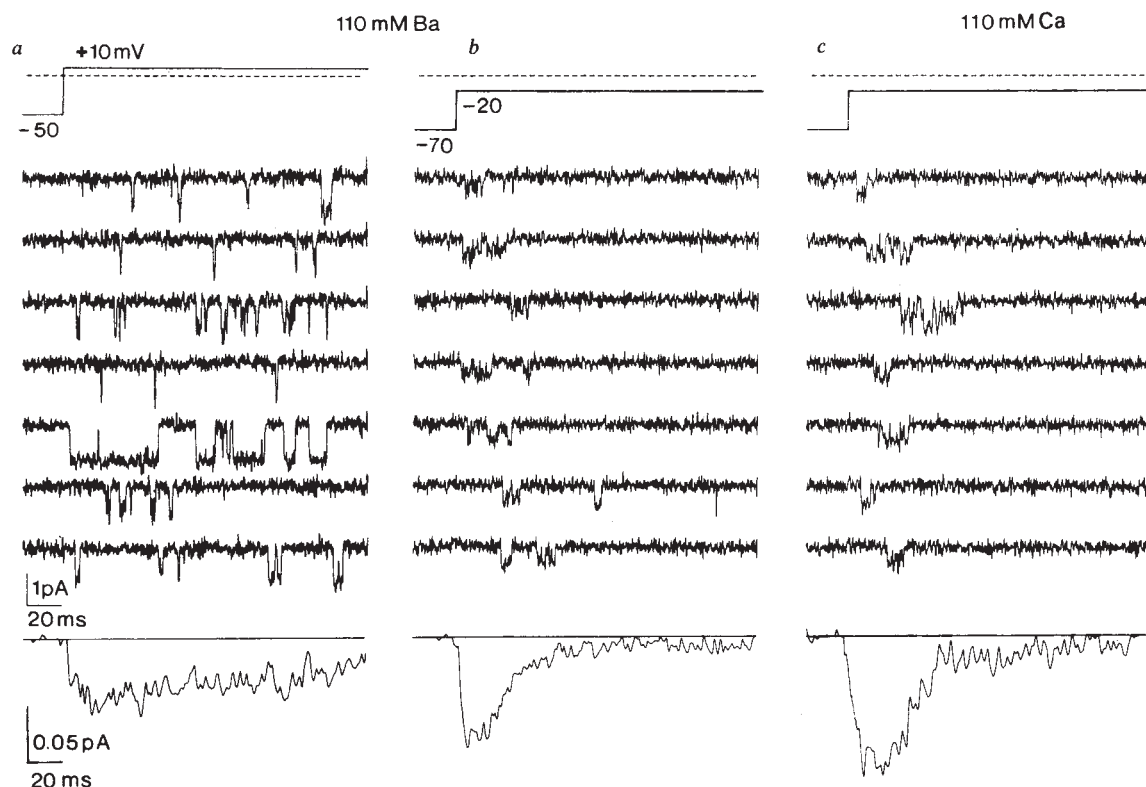


Fig. 1 Evidence for the coexistence of two distinct types of Ca channel in guinea pig ventricular cells. *a, b*, Cell-attached patch recordings with 110 mM Ba in the pipette. Activity of previously described type of Ca channel (*a*) or novel type of Ca channel (*b*) was evoked by appropriate voltage clamp protocols (top row). Zero membrane potential is indicated by the dotted line. A voltage step from a holding potential of -70 mV to a test potential of $+10$ mV produced a mixture of both types of activity (traces not shown). Selected sweeps with clearly detectable channel activity are illustrated. Averaged currents (bottom row) were obtained from all sweeps in each run (294 sweeps in *a*, 270 sweeps in *b*). The amplitude of the averaged currents is not meant to represent the overall balance between the two types of activity. Cell B2E. *c*, Current traces recorded with 110 mM Ca in the pipette with the same voltage protocol as in *b*. The average current (bottom) was obtained from 179 sweeps. Note that the unitary current amplitude and average current time course are very similar with either 110 mM Ca or 110 mM Ba as the charge carrier. Cell B2C.

Methods. Single guinea pig ventricular myocytes were freshly dissociated by enzymatic dispersion²². Cell-attached patch-clamp recordings²³ were made at room temperature (21°C) with depolarizing pulses applied every 3 s. Patch pipettes contained 110 mM BaCl_2 or CaCl_2 plus 10 mM HEPES (pH 7.5 with tetraethylammonium-OH). The membrane potential outside the patch was zeroed by an external solution containing (in mM): K-aspartate 140, EGTA 10, HEPES 10 (titrated to pH 7.5 with KOH). The current traces were filtered (-3 db at 1 kHz, 8-pole Bessel filter), digitized at 5 kHz, stored and analysed on a laboratory computer. Capacity and leak currents were subtracted digitally.

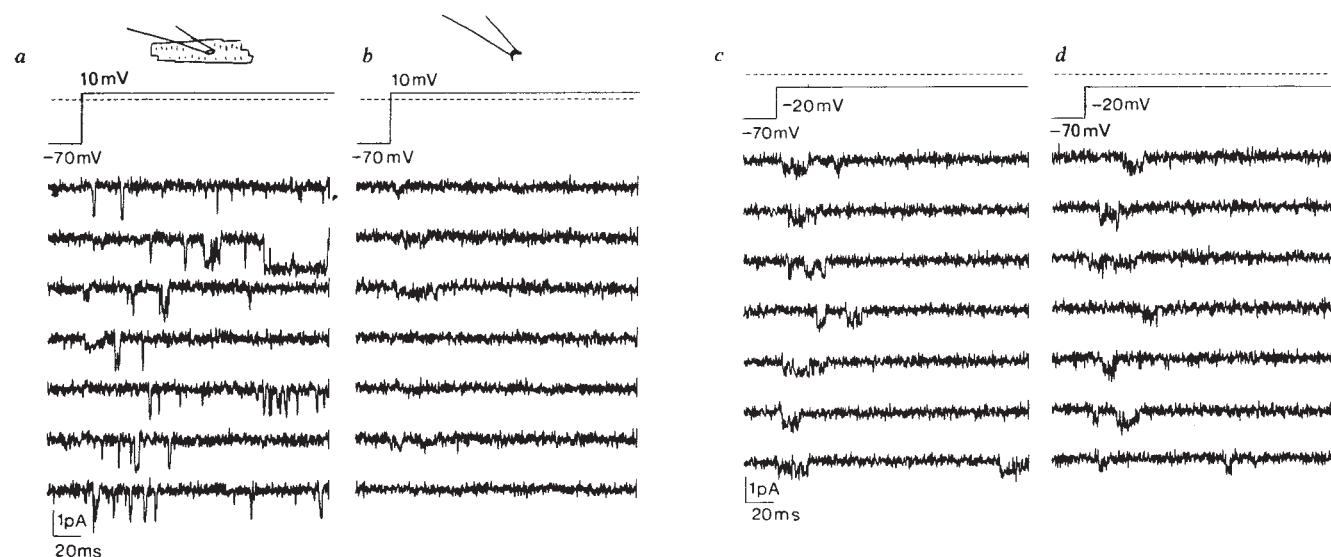


Fig. 2 Differential effect of patch excision on the two types of Ca channels in a single patch. *a, c*, Cell-attached patch recordings of combined L-type and T-type Ca channel activity 10 min before excision (*a*) and exclusively T-type Ca channel activity 6 min before excision (*c*). *b, d*, Excised patch recordings 1 min after excision (*b*) and 6 min after excision (*d*). Activity of the small-conductance T-type Ca channel remains unchanged (*b, d*), while activity of the large-conductance L-type channel could no longer be detected (*b*). Cell B2E.

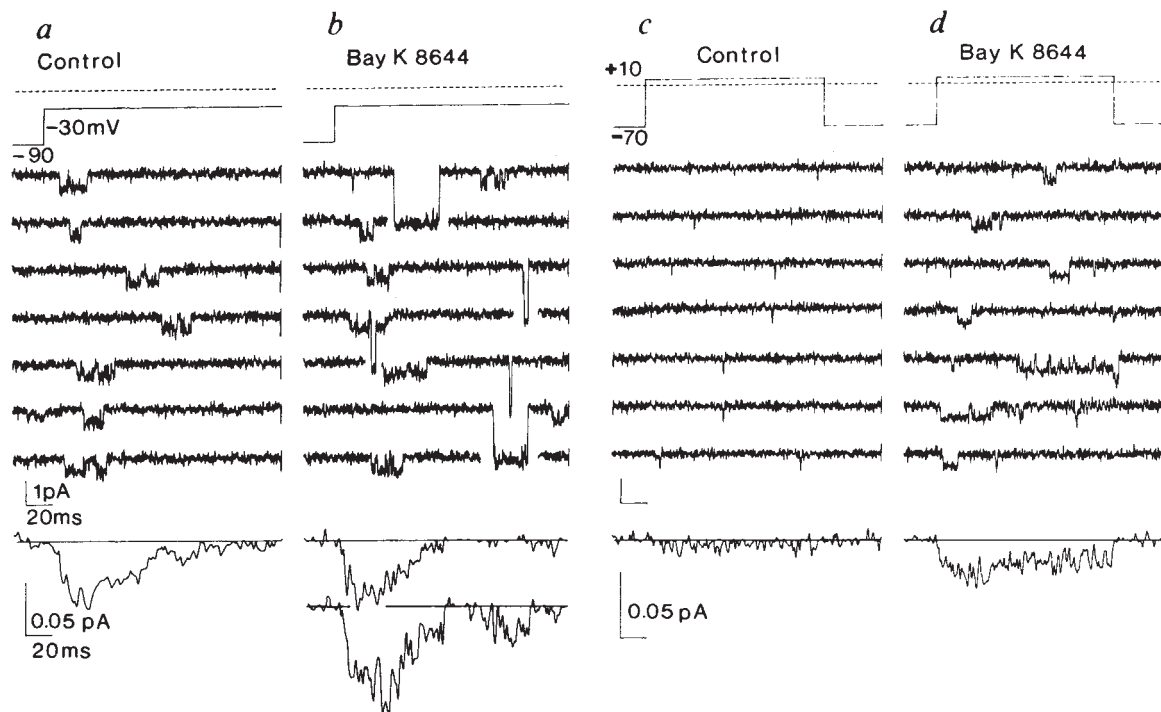


Fig. 3 Differential effect of the dihydropyridine Ca channel agonist Bay K 8644 on the two channel types with Ba or Ca as the charge carrier. *a, b*, Current traces with 110 mM Ba and 33 μ M tetrodotoxin in the pipette from a patch containing both channel types, before (*a*) and after (*b*) exposure of the cell to 5 μ M Bay K 8644. Sweeps with no detectable openings not shown. Lowest traces in each column are averaged currents from all sweeps (234 in *a*, 274 in *b*). Column *b* also includes a smaller average current signal (just above lowest trace), obtained from 194 selected sweeps that showed no detectable activity of the large conductance channel. Cell B3D. *c, d*, 110 mM Ca in the pipette. Current traces from a patch with only L-type Ca channel activity before (*c*) and after (*d*) exposure of the cell to 5 μ M Bay K 8644. Sweeps with no detectable openings not shown. Mean currents (below individual traces) are averages of 151 (*c*) and 518 (*d*) sweeps. Cell B7E.

33 μ M; see Fig. 3) or lidocaine (0.5 mM) known to abolish Na channel current; they are about 50 times larger than predicted from estimates of the permeability ratio P_{Ca}/P_{Na} for the Na channel⁹ and measurements of unitary Na currents^{10,11}. Furthermore, when studied with pipettes of comparable dimensions (~ 2 –2.5 μ m diameter), the number of functional Na channels with 130 mM Na as charge carrier is typically greater than 10 per patch, while the number of the new type of Ca channels with 110 mM Ba or Ca ranged from 0 to at most 3. The Ca channel activity in Fig. 1*b, c* can be labelled 'T' for its markedly transient time course, very much like the T-type Ca channel in dorsal root ganglion (DRG) neurones⁸; by way of contrast, the previously described type of cardiac Ca channel (Fig. 1*a*) is large in unitary conductance (with Ba as the charge carrier) and long-lasting in time course, and will be referred to as 'L'.

One striking difference between T-type and L-type Ca channels in heart is seen when patches are excised (Fig. 2). As little as 1 min after manual excision of the patch from the cell-attached to the inside-out patch configuration, the activity of the L-type Ca channel (Fig. 2*a*) disappeared irreversibly (*b*), as previously reported², whereas T-channel openings (*c*) could still be recorded many minutes after the excision (*d*), without a clear sign of decreasing activity. This contrast between the two channel types was seen consistently in each of five patches, with T-channel activity sometimes remaining up to the end of recordings as long as 20 min.

Organic Ca channel antagonists are widely used in the management of cardiovascular disorders¹² and in attempts at isolation and purification of Ca channels¹³, so it is important to look for differences in the responsiveness of the two types of Ca channel. We found that the dihydropyridine Ca channel antagonist nimodipine (5 μ M) halved the average current through the L-type Ca channel with 110 mM Ba as the charge carrier, whereas at comparable holding potentials (-70 mV), the same dose of nimodipine failed to reduce T-channel activity

in four patches with either 110 mM Ca or Ba as the charge carrier. The contrast between Ca channel types was also found for the dihydropyridine Ca agonist Bay K 8644 (ref. 14). Figure 3*a, b* shows unitary currents from a patch that contained both channel types, with 110 mM Ba as the charge carrier. At a test potential of -30 mV in the absence of the drug (panel *a*), activity consists solely of T-type channel openings, as expected from the differing voltage sensitivity of the two channel types. After addition of 5 μ M Bay K 8644 (Fig. 3*b*), depolarizations to the same test potential continue to evoke T-type channel activity, but with the additional appearance of unitary currents of large amplitude (~ 2 pA). The additional events are typical for Bay K 8644-promoted openings of L-type channels^{15–17}, which are known to appear at negative test potentials where L-type channel activation is normally undetectable¹⁵. The lack of effect of Bay K 8644 on T-channel activity was documented by averaging only those sweeps which lacked large-amplitude unitary currents. The resulting mean current signal (next to last trace in Fig. 3*b*) closely matched the amplitude of the mean current before drug addition (Fig. 3*a*). When Ca was the charge carrier, the T-type channel was also insensitive to 5 μ M Bay K 8644 (not shown), whereas L-type channels clearly responded to the drug (Fig. 3*c, d*). The patch of Fig. 3*c* contained only the L-type channel. At $+10$ mV, its activity is apparent as very brief openings of ~ 0.6 pA amplitude. After addition of 5 μ M Bay K 8644, many long-lasting openings are observed and the averaged current is increased ~ 5 -fold (Fig. 3*d*). The open-time histogram after Bay K 8644 was composed of two exponential components corresponding to short-lived and long-lived open states, one with a very short time constant (< 0.3 ms) as in the absence of drug, and an additional component with a time constant ~ 10 times as long. Thus, our experiments indicate that the contrast between Ca-agonist responsive and unresponsive activity arises from different channel types rather than different charge carriers: T-type Ca channels are unresponsive even with Ba as the charge

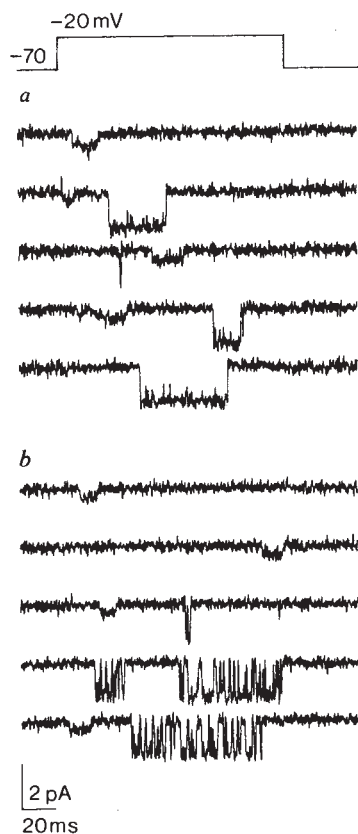


Fig. 4 Different effects of Cd on T-type Ca channels. Selected current traces from two cell-attached patches. Pipettes contained 50 mM Ba as the charge carrier with no Cd (a) or 10 μ M Cd (b). Bath solution contained 5 μ M Bay K 8644. Upper trace shows the voltage clamp protocol. Cells G34C, G34K.

Received 7 February; accepted 22 May 1985.

1. Reuter, H., Stevens, C. F., Tsien, R. W. & Yellen, G. *Nature* **297**, 501-504 (1982).
2. Cavalie, A., Ochi, R., Pelzer, D. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **398**, 284-297 (1983).
3. Brown, A. M., Kunze, D. L. & Yatani, A. *Nature* **311**, 570-572 (1984).
4. Hume, J. R. & Giles, W. J. *gen. Physiol.* **81**, 153-194 (1983).
5. Lee, K. S., Noble, D., Lee, E. & Spindler, A. J. *Proc. R. Soc. B223*, 35-48 (1984).
6. Bean, B. P. *Biophys. J.* **47**, 497a (1985).
7. Carbone, E. & Lux, H. D. *Nature* **310**, 501-502 (1984).
8. Nowycky, M. C., Fox, A. P. & Tsien, R. W. *Biophys. J.* **47**, 67a (1985).
9. Baker, P. F., Hodgkin, A. L. & Ridgway, E. B. *J. Physiol., Lond.* **218**, 709-755 (1971).
10. Sigworth, F. J. & Neher, E. *Nature* **287**, 447-449 (1980).
11. Cachelin, A. B., DePeyer, J. E., Kokubun, S. & Reuter, H. *J. Physiol., Lond.* **340**, 389-401 (1983).
12. Fleckenstein, A. *Calcium Antagonism in Heart and Smooth Muscle* (Wiley, New York, 1983).
13. Campbell, K. P., Strom, M. & Sharp, A. *Biophys. J.* **47**, 514a (1985).
14. Schramm, M., Thomas, G., Towart, R. & Franckowiak, G. *Nature* **303**, 535-537 (1983).
15. Hess, P., Lansman, J. B. & Tsien, R. W. *Nature* **311**, 538-544 (1984).
16. Kokubun, S. & Reuter, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4824-4827 (1984).
17. Ochi, R., Hino, N. & Niimi, Y. *Proc. Jap. Acad. Ser. B60*, 153-156 (1984).
18. Hess, P., Lansman, J. B. & Tsien, R. W. *J. Physiol., Lond.* **358**, 61P (1985).
19. Nowycky, M. C., Fox, A. & Tsien, R. W. *Nature* **316**, 440-443 (1985).
20. Bossu, J. L., Feltz, A. & Thomann, J. M. *Pflügers Arch. ges. Physiol.* (in the press).
21. Yoshii, M., Tsunoo, A. & Narahashi, T. *Biophys. J.* **47**, 433a (1985).
22. Kao, R. L. *et al. Archs Biochem. Biophys.* **203**, 587-599 (1980).
23. Hamill, O., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).

Sarcoplasmic reticulum contains adenine nucleotide-activated calcium channels

Jeffrey S. Smith*, Roberto Coronado†
& Gerhard Meissner*

Departments of * Biochemistry and Nutrition and † Pharmacology, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27514, USA

Rapid calcium efflux from the sarcoplasmic reticulum (SR) is a necessary step in excitation-contraction coupling in skeletal muscle and is thought to be mediated by a calcium channel¹⁻³. Calcium efflux has been studied in fragmented SR vesicles by radioisotope efflux and fluorescence measurements. Several laboratories have reported that adenine nucleotides can stimulate calcium efflux from SR⁴⁻⁷. In recent reports, Ca²⁺ release with a first-order rate constant as high as 100 s⁻¹ has been observed for nucleotide-stimulated Ca²⁺ release from SR vesicles^{8,9}. Also, radioisotope efflux was blocked by Mg²⁺ and micromolar concentrations of the polycationic dye, ruthenium red¹⁻⁷. These high rates of transport are difficult to reconcile with a mechanism other than passive diffusion through a nucleotide-activated 'calcium release channel'⁷⁻⁹. Using the fusion technique for inserting SR proteins into planar lipid bilayers¹⁰, we report here single-channel recordings of calcium release channels from purified 'heavy' SR membranes. Channels have been identified on the basis of their activation by adenine nucleotides, blockade by ruthenium red, and selectivity for divalent cations. Surprisingly, the channel studied here exhibits an unusually large conductance of 170 pS in 50 mM Ba²⁺ while still being capable of discriminating against monovalent cations by a permeability ratio, $P(\text{Ba})/P(\text{Cs}) = 11.4$.

In order to eliminate the contribution from SR monovalent cationic and anionic channels, which comprise the bulk of SR conductance¹¹, and to provide a large divalent cation driving force across the bilayer, heavy SR vesicles were first fused into planar bilayers in asymmetric choline chloride, followed by perfusion with Tris/HEPES (*cis* side) and Ba/HEPES or Ca/HEPES (*trans* ground side). (We chose the sidedness of the buffers based on the previous observations of Miller, which suggest that SR vesicles insert into planar bilayers such that the *cis* chamber corresponds to the myoplasmic space, while the *trans* chamber is equivalent to the lumen of the SR¹².) In our conditions, and according to our sign convention, the Nernst reversal potential $E_{\text{divalent}} = +125$ mV and E_{Tris} is nominally minus infinite, and $E_{\text{HEPES}} = 0$ mV. Thus, divalent cation-selective channels can be easily identified as deflections towards negative current (shown as upward deflections in Fig. 1) that

carrier (Fig. 3a, b) and L-type Ca channels are responsive even if Ca carries current (Fig. 3c, d).

The two types of Ca channels also showed different sensitivities to block by Cd ions (Fig. 4). The presence of micromolar Cd in the patch pipette inhibited L-type Ca channel activity; this was seen most clearly during the long-lasting openings promoted by Bay K 8644 as rapid flickering block¹⁸. In contrast to the clearcut response of L-type channels at 10 μ M Cd, T-type Ca channel openings were not detectably changed in their amplitude or duration. The difference in the Cd responsiveness of T-type and L-type Ca channels was observed in each of four patches containing both channel types. This information from single-channel recordings fits well with results from whole-cell recordings in sensory neurones^{19,20}, and neuroblastoma cells²¹.

We performed whole-cell recordings with 5 mM external Ca or Ba and found an extra component of TTX-resistant inward current with properties corresponding to T-type Ca channels. Its peak amplitude was 5-10% of the peak L-type Ca channel current. Unlike maintained, TTX-resistant inward currents previously described in ventricular cells^{4,5}, T-type current decays rapidly; it shows similarities, however, in its activation at negative potentials and its relative insensitivity to Cd ions⁵. Our results are consistent with the two components of Ca channel current described by Bean⁶ in atrial cells.

In our ventricular cells, T-channel current is much smaller and decays much more quickly than L-channel current, so it probably contributes relatively little to Ca influx during the action potential plateau and contraction. Like their counterparts in neurones^{7,8}, cardiac T-type channels might have greatest significance for pacemaker depolarization and action potential initiation, electrical phenomena that depend critically on small inward currents at relatively negative potentials.

We thank M. C. Nowycky and A. P. Fox for helpful discussions, So Ching Wong for technical assistance, and the USPHS and Miles Pharmaceuticals for support. The work was done at Yale while B.N. was supported by a fellowship from the International Research and Exchanges Board.

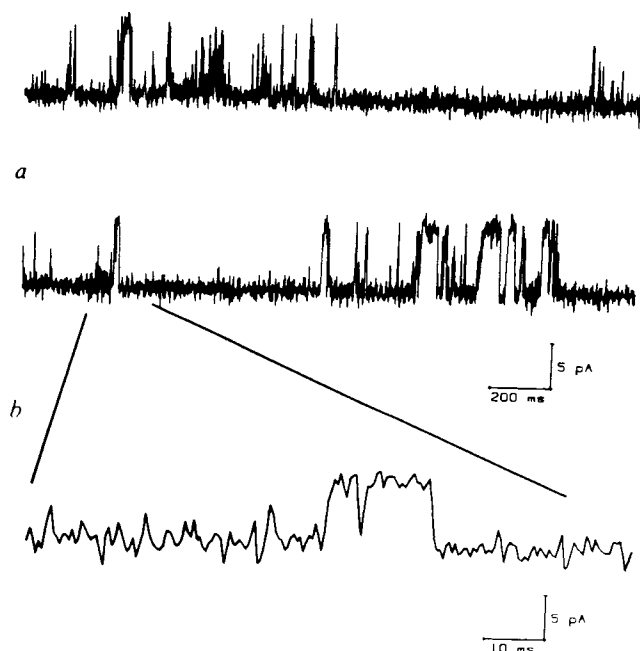


Fig. 1 Incorporation of SR Ca^{2+} channels into planar bilayers. *a*, Single-channel current fluctuations in $2.5 \mu\text{M}$ free Ca^{2+} , 125 mM Tris, 250 mM HEPES pH 7.4 (*cis*) and 50 mM Ba^{2+} , 250 mM HEPES pH 7.4 (*trans*). At -20 mV, the mean open-channel current was 7.55 pA in 50 mM Ba^{2+} . *b*, Single-channel fluctuations from *a* on an expanded time scale.

Methods. 'Heavy' membrane vesicles were prepared as described in ref. 7 and stored at -65°C until use. Planar lipid bilayers containing phosphatidylethanolamine (25 mg ml^{-1} , bovine brain; Avanti Polar Lipids) and phosphatidylserine (25 mg ml^{-1} , bovine brain; Avanti) in decane were painted across a 400- μm hole in a Lexan cup inserted into a cut-away polyvinyl chloride block. In all experiments the *cis* bilayer chamber is defined as the side to which SR vesicles are added; the opposite side is referred to as the *trans* chamber (held at virtual ground). Addition of quick-thawed membrane vesicles to the *cis* bilayer chamber resulted in step-like fusion events. The macroscopic reversal potential in choline chloride (0.25 M *cis*/50 mM *trans*), 5 mM CaCl_2 , 5 mM Tris-HEPES pH 7.4, was $+35$ mV, indicating a high degree of anion specificity. After vesicle-bilayer fusion, both chambers were extensively perfused with chloride-free buffer to reduce the contribution of background chloride-specific conductance. The *cis* chamber was perfused with 125 mM Tris, or 80 mM Cs^+ , 250 mM HEPES, 100 μM Ca-EGTA ($2.5 \mu\text{M}$ free Ca^{2+} ; ref. 20) pH 7.4 and the *trans* chamber with 50 mM Ca^{2+} or Ba^{2+} , 250 mM HEPES pH 7.4. All salts of HEPES were prepared by starting with 250 mM HEPES free acid and titrating to pH 7.4 with the corresponding base (Tris) or metal hydroxide (Ca, Ba or Cs). Current traces in all figures are oriented so that open-channel events are directed upwards (towards negative current) and represent divalent cation-mediated current from the *trans* side to the *cis* side. All experiments were recorded at room temperature (20°C). Records were taken on FM tape, filtered at 0.3 kHz (-3 dB point from 8-pole Bessel filter) and digitized at 2.0 kHz for computer analysis. The duration of open events was measured using computer programs with two threshold detectors placed between the baseline and open peak current. One detector (discriminator (discr.)) was placed at 1 standard deviation (s.d.) from the mean baseline current and the second detector (discr. 2), at 1 s.d. from the mean single-channel unitary current. Open events are defined as transitions that cross discr. 1 and discr. 2 and remain above discr. 2 for two or more consecutive points. Open durations were sorted into 100 bins (2 ms per bin) and fitted by eye to the sum of one or two exponential terms. Events < 2 ms were not included in the fit.

revert at positive potentials. Events of this kind are shown in Fig. 1*a, b*. By using rabbit skeletal muscle SR vesicles derived from heavy terminal cisternae and recovered from the 36–45% region of sucrose density gradients, these channels have been recorded in 90% of all fusion trials ($n > 25$). This fraction was chosen as it is the most active in releasing Ca^{2+} in isotope flux experiments⁷. 'Light' (30–32% region of sucrose density

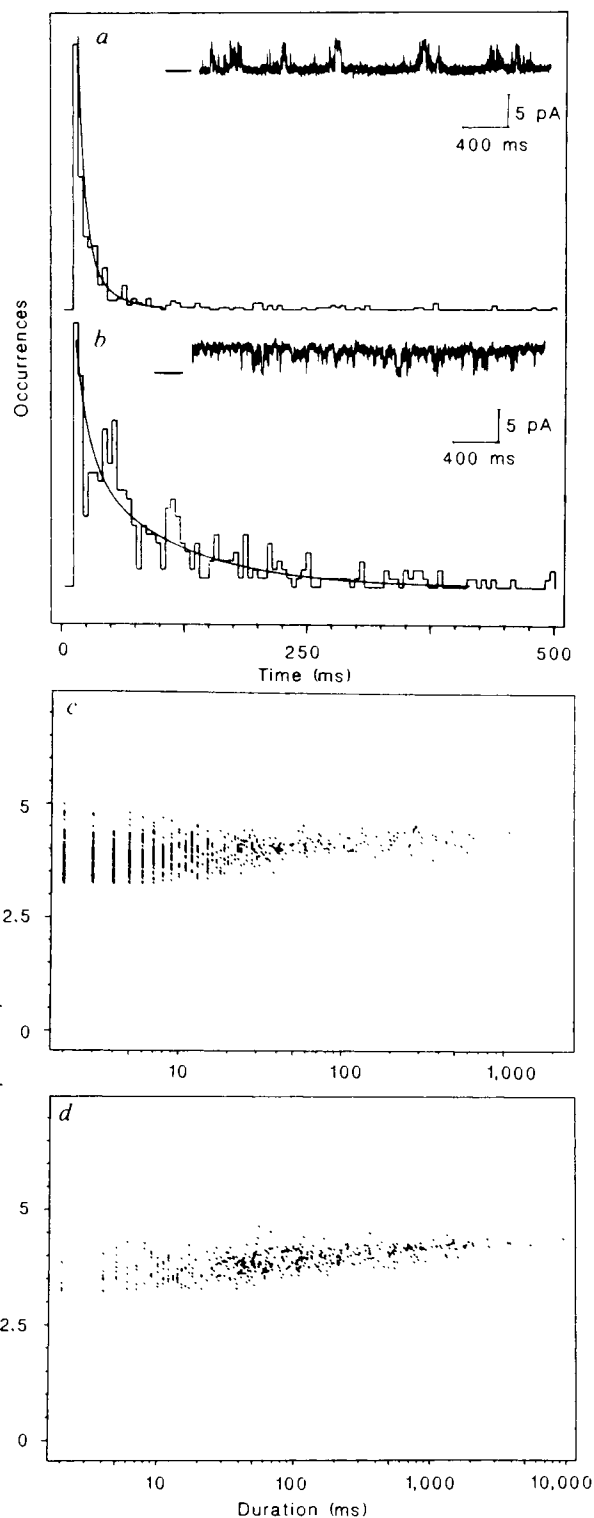


Fig. 2 Nucleotide stimulation of the SR Ca^{2+} channel. *a*, Open-time histogram of data recorded at 0 mV in 125 mM Tris, 250 mM HEPES, 100 μM Ca-EGTA, pH 7.4 (*cis*) and 50 mM Ca^{2+} , 250 mM HEPES pH 7.4 (*trans*). Inset, a representative current trace from the data used to plot the histogram. The horizontal bar to the left of the current trace denotes the closed-channel current. The data were fitted to the sum of two exponentials with the parameters: $\tau_1 = 8.4$ ms, y-intercept 1 = 459, $\tau_2 = 24.4$ ms, y-intercept 2 = 60. *b*, Open-time histogram of data recorded at 0 mV in Tris/Ca as in *a*, with 1 mM Tris-ATP added (*cis*). Inset, a representative current trace in the presence of 1 mM ATP. Data were fitted as in *a* with the parameters: $\tau_1 = 13.5$ ms, y-intercept 1 = 29, $\tau_2 = 95.3$ ms, y-intercept 2 = 16. *c*, Plot of open-event amplitude versus duration for data in *a*. Total open events = 881. Mean open-channel current = 4.2 pA. *d*, Plot of open-event amplitude versus duration for data in *b*. Total of open events = 460. Mean open-channel current = 4.14 pA.

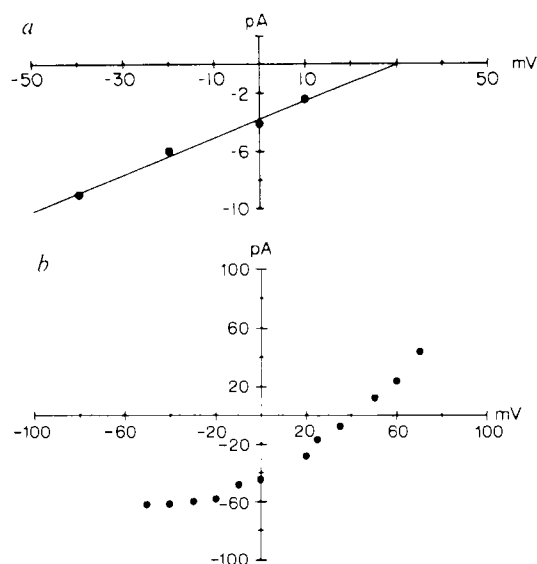


Fig. 3 Single-channel and macroscopic current-voltage (I - V) relationship for nucleotide-stimulated SR Ca^{2+} channels. *a*, Single-channel (I - V) plot for Ca^{2+} channels in 125 mM Tris, 250 mM HEPES, 100 μM Ca-EGTA, 1 mM ATP pH 7.4 (*cis*) and 50 mM Ca^{2+} , 250 mM HEPES pH 7.4 (*trans*). A linear regression fit of the data points gives a reversal potential of +30 mV and a slope conductance of 125 pS. *b*, Macroscopic (I - V) plot for channels in 80 mM Cs^+ , 250 mM HEPES, 100 μM Ca-EGTA, 1 mM ATP pH 7.4 (*cis*) and 50 mM Ba^{2+} , 250 mM HEPES pH 7.4 (*trans*). $P(2+)/P(+)=8.74$ in *a* and 11.4 in *b*, calculated using the expression: $P(2+)/P(+)=\frac{(+)/4(2+)}{[(1+e^{-FV/RT})]/e^{(FV/RT)}}$, assuming no HEPES permeability and that activity is equivalent to the indicated concentration.

gradients) SR vesicle fractions, thought to be derived from the longitudinal cisternae, were almost totally devoid of the channels. This finding is in agreement with ref. 7 in which 'control' vesicles lacking significant Ca^{2+} release activity were obtained from the 30–32% region of sucrose density gradients. The single-channel slope conductance with Ba^{2+} as the current carrier, calculated for channels incorporated into bilayers as shown in Fig. 1, was found to be 170 pS. By expanding the time scale, as in Fig. 1*b*, it is shown that many open-channel events occur on a millisecond time scale.

Figure 2 shows open-time histograms constructed from channel fluctuations recorded with Ca^{2+} as the current carrier at 0 mV before (*a*) and after (*b*) addition of 1 mM ATP *cis*. Representative current traces are shown for each histogram. In Fig. 2*a*, the data were fitted to the sum of two exponentials with values $\tau_1=8.4$ ms and $\tau_2=24.2$ ms. The presence of two open-state lifetimes, as suggested by the fit, could also be anticipated from the current traces of Fig. 1, where both long and short open-channel fluctuations are discernible. In Fig. 2*b*, where records were taken after addition of 1 mM ATP *cis*, the fit gave time constants of $\tau_1=13.5$ ms and $\tau_2=95.3$ ms. Thus, one effect of nucleotide is to increase the duration of the long open state. From the same data it was found that ATP also increased the fraction of channel open time fourfold, from 0.25 to 0.95; this increase in channel open time is due largely to an increase in the fraction of open events having the long lifetime, which increased threefold, from 0.12 to 0.36. Figure 2*c*, *d* shows open-event amplitude/duration plots before (*c*) and after (*d*) addition of 1 mM ATP *cis*. The average open-channel current for both long and short events in the absence or presence of ATP was found to be 4.17 ± 0.032 pA. Thus, the open-channel current does not seem to be affected by ATP. ATP did not seem to exert its activating effect through a protein kinase or some other ATP hydrolysing reaction as the same activating effect was observed when the non-hydrolysable ATP analogue AMP-PCP was used. The physiological significance of channel activation by adenine nucleotides, which are present in muscle at millimolar levels¹³,

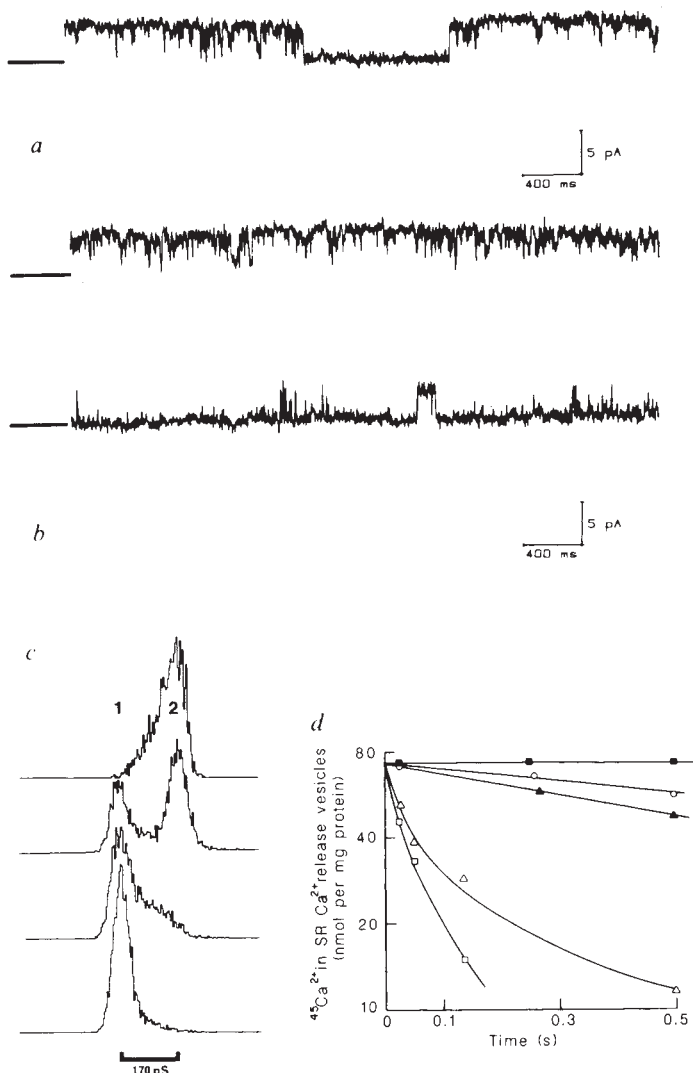
is presently unclear. As the levels of nucleotide are thought to remain constant during excitation-contraction coupling¹³, it is unlikely that ATP functions as a second messenger. Therefore a mechanism must exist in intact muscle by which constant channel activation by nucleotides is prevented. Results from ^{45}Ca vesicle flux experiments (see Fig. 4*d* and refs 7, 9) indicate that Mg^{2+} is an efficient blocker of nucleotide-induced calcium release from SR vesicles. Considering that free Mg^{2+} in muscle is thought to be present at concentrations between 0.2 and several millimolar^{14,15}, this interaction with Mg^{2+} might be one mechanism that acts *in vivo* to prevent constant channel activation by cytoplasmic ATP.

The voltage dependence and selectivity of the nucleotide-stimulated SR calcium channel were investigated in the experiments shown in Fig. 3. Figure 3*a* is a plot of single-channel current against voltage, obtained from an experiment in which SR calcium channels were activated with 1 mM ATP *cis* in Tris/Ca-HEPES, pH 7.4. Single-channel current transitions could not be resolved at voltages greater than +10 mV as in the presence of millimolar nucleotide concentrations, the channels are almost always found in the open state (see Fig. 2*b*). A slope conductance of 125 pS and a reversal potential of +30 mV were obtained by linear extrapolation of the data points. Using an expression derived from constant field electrodiffusion theory, which considers both divalent and monovalent cation permeability¹⁶, we found $P(\text{Ca})/P(\text{Tris})=8.74$. A genuine reversal potential could be measured, however, for the monovalent-divalent pair $\text{Cs}^+/\text{Ba}^{2+}$. Figure 3*b* is a nucleotide-stimulated macroscopic current with 50 mM Ba^{2+} as the *trans* current carrier and 80 mM Cs^+ as the *cis* current carrier. Caesium was chosen as the *cis* monovalent specifically because it is not conducted to any great degree by SR K^+ channels¹⁷, and we suspected that it might be slightly more permeant than Tris^+ in the SR Ca channel. The current in this experiment reverts at +40 mV, which results in a permeability ratio $P(\text{Ba})/P(\text{Cs})=11.4$. Figure 3*b* shows that the barium current is attenuated at negative voltages; this attenuation seems to be due to a decrease in open time and not to a decrease in single-channel conductance (see Fig. 3*a*). In the experiment shown in Fig. 3*a*, the fraction of open time decreased from 0.95 at 0 mV to 0.40 at -40 mV. The voltage dependence of the channel may have important ramifications concerning the mechanism of calcium release from SR.

In agreement with flux experiments in vesicles^{7,9}, the nucleotide-activated channel in bilayers is inhibited by micromolar concentrations of the polycationic dye ruthenium red. Figure 4*a* is a control current trace from an experiment in which a SR calcium channel has been activated by 1 mM ATP *cis* at 0 mV. This trace was selected because it contains a rare, long closed channel event. Addition of 1 μM ruthenium red (Fig. 4*b*) reduced the fraction of open time from 0.95 (control) to <0.02. Figure 4*c* illustrates the time course of block by ruthenium red. Current histograms were made from the experiment in Fig. 4*a*, *b* over a period of 12 s, following addition of 1 μM ruthenium red *cis*. Peaks labelled 1 and 2 represent the most frequent current values for each record and correspond to the closed- and open-channel current levels, respectively. The open-channel peak disappears with time. Thus, ruthenium red appears to affect the SR calcium channel by decreasing the probability of open-channel events. A second effect of ruthenium red is a decrease in the frequency of events that are of the long open type. The block produced by ruthenium red was irreversible in that extensive perfusion did not overcome the blockade.

The similar gating behaviour of the two SR Ca^{2+} conducting pathways, seen in the planar bilayer and the previous vesicle studies⁷⁻⁹, suggests that they both represent events mediated by the same ion channel. We tested this assumption by measuring $^{45}\text{Ca}^{2+}$ efflux from heavy SR vesicles using a radioisotope efflux/rapid quench assay. The SR membrane fraction used in our bilayer experiments was passively loaded with 5 mM $^{45}\text{Ca}^{2+}$ and subjected to conditions identical to those in the *cis* chamber during a bilayer experiment. Figure 4*d* shows that the vesicles

Fig. 4 Inhibition of the nucleotide-stimulated SR Ca^{2+} channel by ruthenium red. *a*, SR Ca^{2+} channel in 50 mM Ca^{2+} , 250 mM HEPES pH 7.4 (*trans*) and 125 mM Tris, 250 mM HEPES, 2.5 μM free Ca^{2+} , 1 mM ATP pH 7.4 (*cis*) at 0 mV. The horizontal bar to the left of the trace denotes the closed-channel current. *b*, Same experiment as *a*, 12 s after addition of 1 μM ruthenium red *cis*. *c*, Time course of ruthenium red block. Current histograms were obtained at 0 mV by analysing 4-s segments of current records (sampling rate 2 kHz) in a pulse height analyser. The histograms were constructed by placing the sample current in the bin corresponding to that particular current value. Thus, the peaks represent the most frequent values of conductance for a given record. The histograms, from top to bottom, represent current before addition of ruthenium red, and 0–4, 4–8 and 8–12 s after addition of 1 μM ruthenium red *cis*. *d*, Measurement of $^{45}\text{Ca}^{2+}$ efflux from passively loaded SR vesicles using a rapid quench technique⁹. Release of $^{45}\text{Ca}^{2+}$ was initiated by diluting vesicles passively loaded with 5 mM $^{45}\text{Ca}^{2+}$ into 125 mM Tris, 250 mM HEPES pH 7.4 medium containing amounts of EGTA, Ca^{2+} , Mg^{2+} , AMP-PCP and ruthenium red so that the release media after the dilution of the vesicles contained 10 mM Mg gluconate plus 10 μM ruthenium red (■), 5 mM EGTA plus 4.94 mM Ca-HEPES (5 μM free Ca^{2+}), in the absence (○) or presence (□) of 1 mM AMP-PCP, 1 mM AMP-PCP plus 1 μM ruthenium red (△), or 1 mM AMP-PCP plus 2.5 mM Mg gluconate (▲). $^{45}\text{Ca}^{2+}$ release was blocked at times from 25 to 500 ms by the addition of 10 mM Mg^{2+} and 10 μM ruthenium red and the amounts of $^{45}\text{Ca}^{2+}$ remaining with SR Ca^{2+} release vesicles were determined by Millipore filtration (ref. 7).



retained ~ 75 nmol $^{45}\text{Ca}^{2+}$ per mg protein when diluted into a solution containing the two calcium efflux blockers Mg^{2+} and ruthenium red^{1–7}. Dilution into a solution containing 5 μM free Ca^{2+} (○) resulted in an initial release rate of 40 nmol $^{45}\text{Ca}^{2+}$ per mg protein per s. In the presence of 1 mM AMP-PCP (□), the vesicles released $^{45}\text{Ca}^{2+}$ with an initial rate of $\sim 1,000$ nmol $^{45}\text{Ca}^{2+}$ per mg protein per s. Ruthenium red at 1 μM (△) reduced the efflux rate by about twofold, indicating a significant block of the Ca^{2+} efflux pathway. In other studies not shown here, we found that inhibition of $^{45}\text{Ca}^{2+}$ release was dependent on the time of exposure to 1 μM ruthenium red, as has been noted in the bilayer studies (see Fig. 4c). Dilution into a medium containing 2.5 mM Mg^{2+} in addition to 1 mM AMP-PCP (▲) resulted in a 15-fold reduction in $^{45}\text{Ca}^{2+}$ release rate. Ruthenium red at 10 μM in combination with 10 mM Mg^{2+} (■) produced a much more rapid and complete inhibition of Ca^{2+} efflux from SR vesicles. We plan to study more closely this effect of Mg^{2+} on the SR Ca^{2+} channel in bilayers, as we believe the effects of

Mg^{2+} , nucleotides and Ca^{2+} are important in the regulation of the SR Ca^{2+} channel *in vivo*.

We have demonstrated here that calcium channels from terminal cisternae-derived heavy SR vesicles can be incorporated into planar lipid bilayers. The channels are activated by adenine nucleotides and inhibited by ruthenium red. The channels also exhibit an unusually large conductance (170 pS in 50 mM Ba^{2+}) which is 10–20 times the conductance measured for brain or transverse tubule membrane calcium channels incorporated in planar bilayers^{18,19}. This very large conductance, together with the rapid kinetics seen in both bilayers and vesicle studies, are exactly the attributes one might expect for a calcium channel that can mediate large ion fluxes on a millisecond time scale. These two factors lead us to propose that the calcium channel studied here is the putative calcium release channel in SR.

This work was supported by NIH grants AM 18687 and GM 32824. We thank Dr H. Affolter for the use of his computer program.

Received 4 February; accepted 3 June 1985.

1. Ebashi, A. *Rev. Physiol.* **38**, 293–313 (1976).
2. Endo, M. *Physiol. Rev.* **57**, 71–108 (1977).
3. Bianchi, C. P. & Frank, G. B. *Can. J. Physiol. Pharmacol.* **60**, 415–588 (1982).
4. Ogawa, Y. & Ebashi, S. *J. Biochem., Tokyo* **80**, 1149–1157 (1976).
5. Endo, M. & Kitazawa, O. *Proc. Japan Acad.* **52**, 599 (1976).
6. Morii, H. & Tonomura, Y. *J. Biochem., Tokyo* **93**, 1271–1285 (1983).
7. Meissner, G. *J. Biol. Chem.* **259**, 2365–2374 (1984).
8. Nagasaki, K. & Kasai, M. *J. Biochem., Tokyo* **94**, 1101–1109 (1983).
9. Meissner, G., Darling, E. & Eveleth, J. *Biochemistry* (in the press).
10. Miller, C. & Racker, E. *J. Membrane Biol.* **30**, 283–300 (1976).

11. Meissner, G. *Molec. cell. Biochem.* **55**, 65–82 (1983).
12. Miller, C. *J. Membrane Biol.* **40**, 1–23 (1978).
13. Kushmerick, M. J. in *Handbook of Physiology* Section 10 (eds Peachey, L. D., Adrian, R. H. & Geiger, S. R.) 189–236 (American Physiological Society, Bethesda, Maryland, 1983).
14. Baylor, S. M., Chandler, W. K. & Marshall, M. W. *J. Physiol., Lond.* **344**, 625–666 (1983).
15. Gupta, R. K. & Moore, R. D. *J. Biol. Chem.* **255**, 3987–3993 (1980).
16. Lee, K. S. & Tsien, R. W. *J. Physiol., Lond.* **354**, 253–272 (1984).
17. Coronado, R., Rosenberg, R. & Miller, C. *J. gen. Physiol.* **76**, 425–446 (1980).
18. Nelson, M. T., French, R. J. & Krueger, B. K. *Nature* **308**, 77–80 (1984).
19. Affolter, H. & Coronado, R. *Biophys. J.* (in the press).
20. Fabiato, A. *J. gen. Physiol.* **78**, 457–497 (1981).

Genes for the major protein antigens of the leprosy parasite *Mycobacterium leprae*

Richard A. Young*, Vijay Mehra*||, Douglas Sweetser*, Thomas Buchanan†, Josephine Clark-Curtiss‡, Ronald W. Davis§ & Barry R. Bloom||

* Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Immunology Research Laboratory, Pacific Medical Center, and Departments of Pathobiology and Medicine, University of Washington, Seattle, 98144, USA

‡ Department of Biology, Washington University, St Louis, Missouri 63130, USA

§ Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA

|| Department of Microbiology and Immunology, and Irvington House Institute for Medical Research, Albert Einstein College of Medicine, Bronx, New York 10461, USA

Leprosy, a chronic infectious disease afflicting between 10 and 15 million people, is caused by the obligate intracellular parasite *Mycobacterium leprae*¹. Although *M. leprae* was the first identified bacterial pathogen of man², basic biochemical, immunological, diagnostic and therapeutic investigations have been severely limited because it remains one of the few human pathogens that have not been cultured *in vitro*. An *M. leprae* recombinant DNA expression library was constructed to provide a source of genes encoding proteins relevant for such studies. Monoclonal antibodies directed against *M. leprae* specific antigens³⁻⁷ have been used to isolate the genes encoding the five most immunogenic protein antigens of the leprosy bacillus. We report here that *M. leprae* specific epitopes recognized by all of 13 monoclonal antibodies tested were produced by recombinant phage in *Escherichia coli*.

A recombinant DNA expression library of *M. leprae* DNA was constructed using λ gt11, a bacteriophage vector that is capable of driving the expression of foreign insert DNA with *E. coli* transcription and translation signals⁸⁻¹⁰. An attempt was made to increase the likelihood that all possible foreign coding sequences would be expressed in *E. coli* by using an approach that proved successful in the isolation of *Mycobacterium tuberculosis* genes¹¹. *M. leprae* was purified from an armadillo that had been inoculated with bacillus from a single human patient¹²; DNA was purified from the bacillus¹³ and was mechanically sheared to produce fragments of 1-7 kilobases (kb). *Eco*RI linkers were added to the ends of the DNA fragments to allow insertion at the unique *Eco*RI site of λ gt11. The ligated recombinant DNA was packaged into phage heads and this material was used to infect *E. coli* cells¹⁴. The aim of this approach was to generate DNA fragments with random endpoints throughout the foreign genome and to then produce recombinant phage in sufficient numbers that insert endpoints occurred at each base pair in the pathogen genome. This strategy should ensure that all coding sequences are inserted in the correct transcriptional orientation and translational frame to be expressed as a fusion protein with the β -galactosidase encoded in λ gt11.

The *M. leprae* DNA library constructed in this manner contained 2.5×10^6 individual recombinant phage. This library was amplified in *E. coli* Y1088 (ref. 9) by producing a plate stock¹⁴ whose titre was 2×10^{11} PFU (plaque-forming units) ml⁻¹. The amplified library consisted of 25% recombinants whose foreign DNA insert lengths averaged 2 kb, as determined by DNA restriction endonuclease analysis of 25 independent phage clones. As the *M. leprae* genome consists of $\sim 10^6$ base pairs (bp)¹³, it is likely that this library comprehensively represents the DNA of the bacillus.

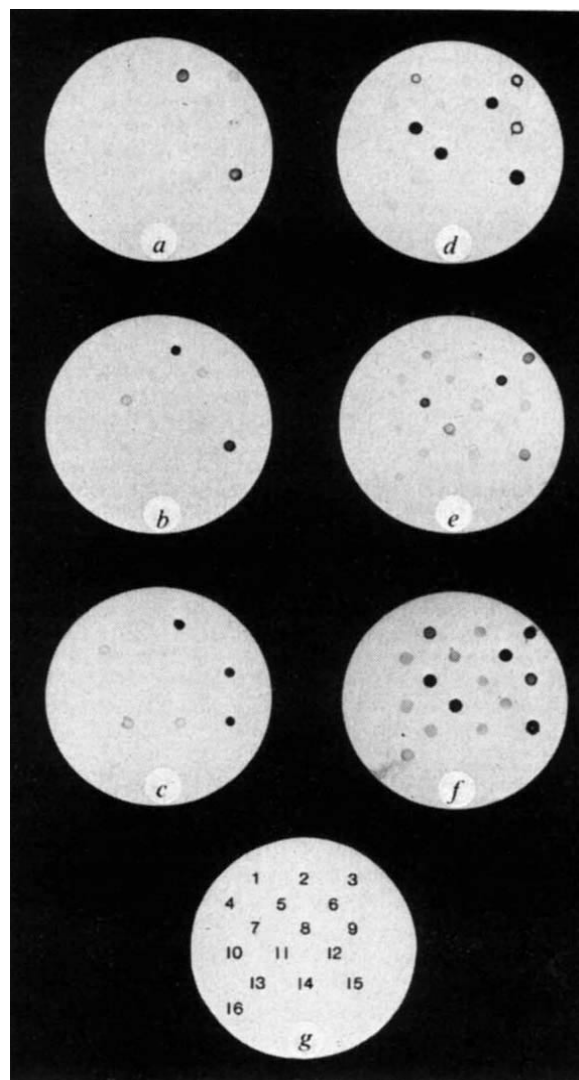


Fig. 1 Arrays of antigen from *M. leprae* recombinant DNA clones probed with individual monoclonal antibodies. Drops containing $\sim 10^4$ PFU each of 15 cloned λ gt11 recombinants were arrayed on lawns of *E. coli* Y1090. The phage were grown and the antigens blotted and probed with individual monoclonal antibodies at $\sim 1:200$ dilution as described in ref. 11. The monoclonal antibodies used here are: a, MLIIC8; b, MLIIC8; c, MLIIE9; d, MLIH9; e, Y1-2; and f, Cl-1. The recombinant DNA clones are coded in g: 1, Y3159; 2, Y3160; 3, Y3161; 4, Y3162; 5, Y3165; 6, Y3166; 7, Y3170; 8, Y3171; 9, Y3172; 10, Y3173; 11, Y3174; 12, Y3175; 13, Y3176; 14, Y3177; 15, Y3178; 16, λ gt11 (this phage plaque produced the background signal for each monoclonal antibody).

The monoclonal antibodies were produced in mice immunized with intact or crude extracts of armadillo-derived, purified *M. leprae*³⁻⁷, and the sizes of the antigens to which they bind are listed in Table 1. All of the antibodies directed against the antigen of relative molecular mass (M_r) 65,000 (65K) were pooled at approximately 1:200 dilution and used to probe 10^6 plaques (0.25×10^6 recombinant plaques) according to protocols described previously^{10,11}. Seventeen plaques produced signals; 15 of these were successfully purified to homogeneity in one or two successive rescreens with the antibody pool.

Recombinant DNA clones isolated in this manner were then arrayed and probed with each of the monoclonal antibodies individually. Figure 1 shows the results obtained with six of the seven antibodies directed against the 65K antigen. All seven monoclonal antibodies, each directed against a different epitope (T. P. Gillis, R. A. Miller, D. B. Young, S. R. Khanolkar and T.B., personal communication), were capable of recognizing antigen produced in *E. coli*. At least four different signal patterns

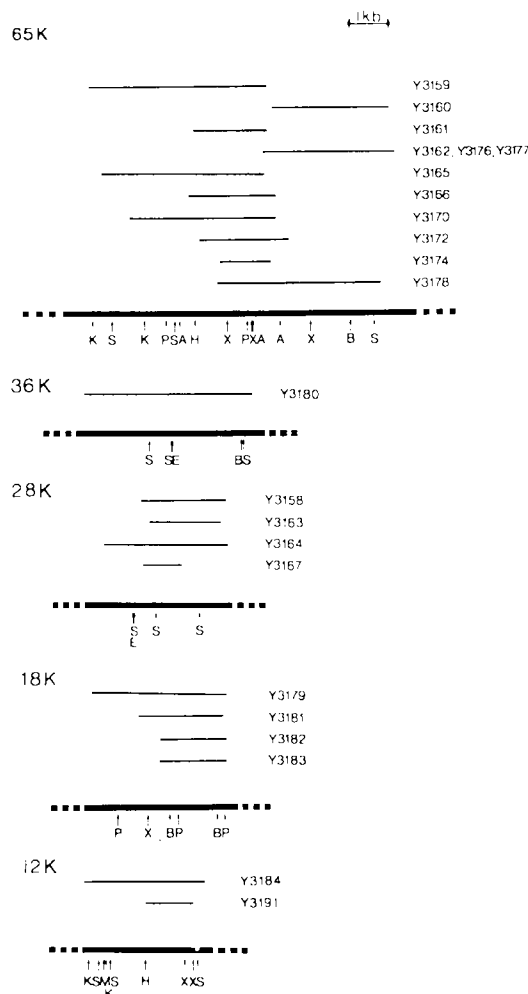


Fig. 2 Restriction maps of *M. leprae* DNA. Recombinant DNA clones isolated by using antibodies directed against the 65K, 36K, 28K, 18K and 12K antigens were subjected to restriction endonuclease mapping. λ DNA was prepared from phage plate stocks according to Davis *et al.*¹⁶. A, *SacI*; B, *BglII*; E, *EcoRI*; H, *HindIII*; K, *KpnI*; M, *BamHI*; P, *PvuI*; S, *Sall*; X, *XhoI*.

were observed in the array of clones. Antibodies MLIIC8, MLIIC8 and MLIIE9 (Fig. 1a, b and c, respectively) produced three distinct patterns. A fourth pattern was generated by antibodies MLIH9, Y1-2 and C1-1 (Fig. 1d, e and f, respectively). Antibody ML-30, directed against an epitope shared with *E. coli*, produced strong signals with all clones, generating a poorly discernible pattern (not shown). There was considerable variation in the number of different epitopes produced by each clone, as evidenced by their reaction with different monoclonal antibodies. While some clones produced antigen that was recognized by only one antibody (for example, clone 13 in Fig. 1c), and one of the clones produced antigen recognized by all seven antibodies (clone 15), most clones produced antigen that was bound by some intermediate fraction of the seven antibodies.

Recombinant DNA clones were also isolated and characterized with antibodies directed against the 36K, 28K, 18K and 12K antigens. Approximately 10^6 λ gt11 plaques were screened with a pool of these monoclonal antibodies. Eleven plaques that produced signals were purified to homogeneity. Clones were arrayed as in Fig. 1 and probed with each of the individual antibodies that comprised the pool. The anti-36K, 28K, 18K and 12K antibodies produced signals with 1, 4, 4 and 2 recombinant clones, respectively.

The insert DNAs of all the recombinant DNA clones isolated using these monoclonal antibodies were mapped with restriction endonucleases. Figure 2 shows the genomic DNA restriction

Table 1 Monoclonal antibodies used to isolate *M. leprae* genes

Antibody	<i>M. leprae</i> antigen	Ref.
MLIIC8	65K	4
MLIIC8	65K	4
Y1-2	65K	7
MLIIE9	65K	4
MLIIE9	65K	4
C1-1	65K	7
ML-30	65K	3
F47 CL9.1	36K	7
SA1.D2D	28K	6
SA1.B11H	28K	6
L7-15	18K	7
ML-06	12K	3

Monoclonal antibodies directed against *M. leprae* protein antigens. The mouse monoclonal antibodies used in this study are tabulated according to the size of the *M. leprae* antigen that they bind and a reference describing their source. All of the antibodies are IgG1.

maps deduced for genes encoding each of the five antigens of interest and illustrates how each of the cloned DNAs aligns with that map. All clones isolated with monoclonal antibodies directed against any single antigen appear to align with a single genomic DNA segment. For example, the insert DNAs of the 12 recombinant clones isolated with the seven different anti-65K monoclonal antibodies overlap sufficiently to allow the construction of a unique genomic DNA restriction map with which all clones align unambiguously. This result indicates that all the anti-65K antibodies recognize epitopes encoded in a single DNA locus, presumably in a single gene. Similarly, the multiple clones isolated with the anti-28K, 18K or 12K antibodies produced overlapping restriction maps.

One concern with the approach used here is that some recombinant clones may be isolated not because they express the protein of interest but because they express an unrelated polypeptide containing a similar or identical immunological determinant. However, when multiple recombinant DNA clones were isolated by using a single monoclonal antibody, all contained overlapping DNA; this suggests that each of the epitopes of interest is encoded by a single genomic DNA segment.

These results attest to the power of the approach used here to detect and isolate specific antigen-coding sequences from λ gt11 recombinant DNA expression libraries. It is striking that all of the monoclonal antibodies used in the present study could recognize antigen produced by recombinant phage in *E. coli*. The nature of the epitopes recognized by these antibodies is presently under investigation. Some of the monoclonal antibodies used here bind small linear peptide determinants. For example, at least two of the anti-65K monoclonal antibodies bind antigen from λ gt11 subclones of Y3178 (Fig. 2) which contain less than 100 bp of insert DNA, encoding at most 33 amino acids (R.A.Y., unpublished data).

In summary, the recombinant DNA strategy used here to express the coding capacity of the *M. leprae* genome in *E. coli* and to detect specific antigenic determinants with monoclonal antibodies, appears to be effective. Each of the genes of interest was isolated by screening no more than 10% of the recombinant DNA library. These results suggest that a similar approach is feasible with genomic DNA from parasites having more complex genomes, such as schistosomes, filaria, leishmania and plasmodia, parasites that cause disease in hundreds of millions of people.

The availability of antigens produced by recombinant DNA technology may make it possible to address some problems in leprosy that cannot be approached by other means. It should be possible to develop simple and specific seroepidemiological assays using these protein antigens, in order to screen populations for individuals producing antibodies to *M. leprae*-specific antigenic determinants. This could make early diagnosis of

leprosy feasible, permitting early treatment to reduce its transmission and prevent the development of nerve damage and deformity.

As resistance to leprosy is provided by cell-mediated immunity¹, the availability of antigens produced by recombinant DNA technology should allow experiments to be performed to address the basic immunological question of whether epitopes recognized by monoclonal antibodies are also recognized by T cells. In addition, a mixture of peptides recognized by helper T cells could, for the first time, provide a specific skin test antigen for use in assessing the immunological status of patients and their contacts—this is urgently needed to permit rapid evaluation of the immunological efficacy of the candidate vaccines against leprosy that are being developed¹⁵. Finally, the identification of antigens involved in cellular immunity could lead to the development of a new generation of vaccines. Because of the intrinsic adjuvant properties of mycobacteria, one attractive approach is to introduce genes whose products provide protection into cultivatable mycobacteria, to produce a vaccine capable of engendering long-lasting cell-mediated immunity.

We thank D. Young, J. Ivanyi, A. H. J. Kolk, W. Britten, T. Gillis and the World Health Organization for monoclonal antibodies, and C. Shepard for the *M. leprae*-infected armadillo. We also thank D. Baltimore, J. Convit, T. Godal, N. Oliver and C. Shepard for advice and discussion and S. Raines for preparing

the manuscript. This research was supported by grants from the WHO/World Bank/UNDP Special Programme for Research and Training in Tropical Diseases; from the Rockefeller Foundation Program for Research in Great Neglected Diseases; from the US National Hansen's Disease Center; and by USPHS grant AI2011. This paper is dedicated to the memory of Dr Charles Shepard.

Received 15 April; accepted 23 May 1985.

1. Bloom, B. R. & Godal, T. *Rev. Infect. Dis.* **5**, 765–780 (1983).
2. Hansen, G. A. *Norsk Mag. Laegevitensk* **4**, 1–88, I-L III (1874).
3. Ivanyi, J., Morris, J. A. & Keen, M. in *Monoclonal Antibodies Against Bacteria* (eds Macario, A. J. L. & Macario, E. C.) (Academic, New York, 1984).
4. Gillis, T. P. & Buchanan, T. M. *Infect. Immunity* **37**, 172 (1982).
5. Coates, A. R. M., Allen, B. W., Hewitt, J., Mitchison, D. A. & Ivanyi, J. *Lancet* **ii**, 167 (1981).
6. Young, D. B., Fohn, M. J., Khanolkar, S. R. & Buchanan, T. M. *Clin. exp. Immun.* (in the press).
7. Engers, H. *et al. Infect. Immunity* **48**, 603–605 (1985).
8. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194–1198 (1983).
9. Young, R. A. & Davis, R. W. *Science* **222**, 778–782 (1983).
10. Young, R. A. & Davis, R. W. in *Genetic Engineering: Principles and Techniques* Vol. 7 (eds Setlow, J. & Hollaender, A.) 29–41 (Plenum, New York, 1985).
11. Young, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 2583–2587 (1985).
12. Draper, P. *Int. J. Lepr.* **44**, 95–98 (1976).
13. Clark-Curtiss, J. E., Jacobs, W. R., Docherty, M. A., Ritchie, L. R. & Curtiss, R. J. *Bact.* **161**, 1093–1102 (1985).
14. Huynh, T., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D.) (IRL, Eynsham, in the press).
15. Bloom, B. R. & Mehra, V. in *New Approaches to Vaccine Development* (eds Bell, R. & Torrigiani, G.) 368–389 (Schwabe & Co., Basel, 1984).
16. Davis, R. W., Botstein, D. & Roth, J. R. *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, New York, 1980).

Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG

R. B. Parekh*, R. A. Dwek*, B. J. Sutton†, D. L. Fernandes*, A. Leung‡, D. Stanworth‡ & T. W. Rademacher*

Departments of *Biochemistry and †Molecular Biophysics, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

‡ Department of Immunology, Rheumatoid and Allergy Research Unit, University of Birmingham, The Medical School, Vincent Drive, Birmingham B15 2TJ, UK

T. Mizuochi§, T. Taniguchi§, K. Matsuta||, F. Takeuchi||, Y. Nagano||, T. Miyamoto|| & A. Kobata§

§ Department of Biochemistry, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108, Japan

|| Department of Medicine and Physical Therapy, Faculty of Medicine, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan

Rheumatoid arthritis (RA) is a widely prevalent (1–3%) chronic systemic disease thought to have an autoimmune component¹; both humoral^{1–4} and cellular^{5,6} mechanisms have been implicated. Primary osteoarthritis (OA) is considered to be distinct from rheumatoid arthritis, and here damage is thought to be secondary to cartilage degeneration. In rheumatoid arthritis, immune complexes are present that consist exclusively of immunoglobulin⁷, implying that this is both the 'antibody' (rheumatoid factor [RF]) and the 'antigen' (most commonly IgG). Autoantigenic reactivity has been localized to the constant-region (C_γ2) domains of IgG^{8,9}. There is no evidence for a polypeptide determinant but carbohydrate changes have been reported¹⁰. We have therefore conducted a study, simultaneously in Oxford and Tokyo, to compare in detail the *N*-glycosylation pattern of serum IgG (Fig. 1) isolated from normal individuals and from patients with either primary osteoarthritis or rheumatoid arthritis. The results, which required an evaluation of the primary sequences of ~1,400 oligosaccharides

from 46 IgG samples, indicate that: (1) IgG isolated from normal individuals, patients with RA and patients with OA contains different distributions of asparagine-linked bi-antennary complex-type oligosaccharide structures, (2) in neither disease is the IgG associated with novel oligosaccharide structures, but the observed differences are due to changes in the relative extent of galactosylation compared with normal individuals. This change results in a 'shift' in the population of IgG molecules towards those carrying complex oligosaccharides, one or both of whose arms terminate in *N*-acetylglucosamine. These two arthritides may therefore be glycosylation diseases, reflecting changes in the intracellular processing, or post-secretory degradation of *N*-linked oligosaccharides.

At least 30 different complex-type bi-antennary oligosaccharides are associated with human serum IgG (Fig. 2). To compare the molar proportions of each of these structures, each serum IgG sample was subjected to controlled hydrazinolysis to release intact its associated oligosaccharide moieties¹¹. Reduction of the reducing terminal *N*-acetylglucosamine residues using NaB³H₄ was then performed to label radioactively each carbohydrate chain. Each labelled oligosaccharide mixture was subjected to exhaustive neuraminidase digestion in order to analyse the distribution of neutral structures. The resulting 'asialo' oligosaccharide mixtures were fractionated by Bio-Gel P-4 (~400 mesh) gel filtration chromatography, a technique that separates neutral oligosaccharides on the basis of their effective hydrodynamic volumes¹² (Figs 2, 3).

A comparison of the individual profiles obtained reveals several interesting points. First, all P-4 chromatograms of asialo mixtures from control individuals (Fig. 3) were essentially identical to one another (data not shown) and similar to those reported previously¹³. As any individual IgG molecule contains, on average, 2.8 *N*-linked oligosaccharides¹⁴, many different IgG subpopulations must exist, each unique with respect to the sequence of its associated oligosaccharides. Further, the overall relative molar contribution of each of the 30 or so structures in the analysis of polyclonal IgG is remarkably constant. The occurrence of this large number of different oligosaccharides is not the result of performing the analysis on polyclonal IgG, because this same 'set' of structures is found on human myeloma proteins¹³ and an analogous situation exists for mouse monoclonal antibodies¹⁴. Second, the P-4 chromatograms of asialo mixtures from patients with rheumatoid arthritis are also essentially the same between patients (data not shown), but differ from control profiles (Fig. 3). Third, the asialo oligosaccharide

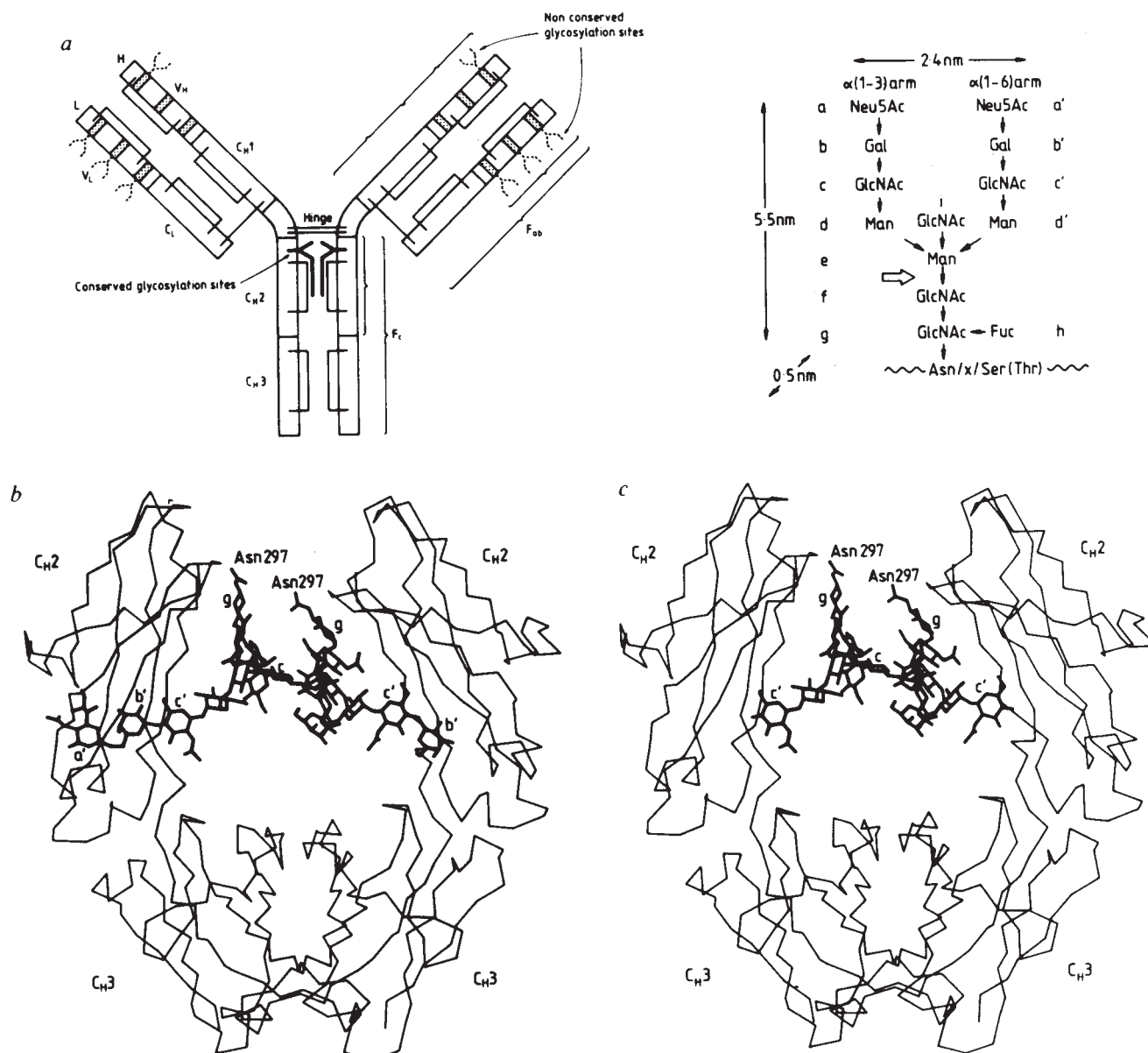


Fig. 1 a. The antibody molecule consists of two heavy (H) and two light (L) chains, linked by disulphide bridges (solid lines) and is divided into homologous regions of sequence (V_H , C_{H1} , C_{H2} and C_{H3}), each of which has an intra-chain disulphide bridge. (The pattern of inter-chain disulphide bridging shown here is characteristic of human subclass IgG1.) In V_H and V_L , the dotted segments represent the hypervariable regions of sequence that, in the three-dimensional structure, together form the antigen-binding site. The conserved asparagine-linked bi-antennary complex oligosaccharide chains are attached to Asn 297 in the C_{H2} domains¹⁶. Oligosaccharide attachment sites are found in the Fab region. Their frequency and location are dependent on the presence of Asn-X-Ser sites in the hypervariable regions^{27,28}. The relative size of an immunoglobulin domain and a fully extended *N*-linked complex oligosaccharide²⁹ are similar. Complex-type oligosaccharides present on IgG can be subdivided into an outer-arm region (a, a', b, b', c, c'), and the core, which is composed of a trimannosyl unit (d, d', e) and a *N,N'*-diacetylchitobiosyl unit (f, g). The 'bisect' GlcNAc (residue i) is linked β (1-4) and the fucose (residue h) is linked α (1-6). The arrow between residues e and f indicates the site of interaction between the two oligosaccharides in b and c. **b.** Refined structure at 2.8 Å of rabbit Fc fragment from the crystal data of Sutton and Phillips (ref. 16 and B. J. Sutton, unpublished data). The two carbohydrate chains, each attached at Asn 297, differ in conformation and may also differ in sequence and bridge the two C_{H2} domains. The α (1-3) arm of the chain (left side) is always devoid of galactose and interacts through its β (1-2)-linked GlcNAc residue (c) with the $\text{Man}\beta$ (1-4) GlcNAc segment of the opposing (right side) oligosaccharide chain (see a). The α (1-3) arm of the right chain extends outwards between the domains with no apparent steric constraints on its length. A Neu5Ac unit (a') is shown on one α (1-6) arm only (left). The electron density for this unit is weak, and experimentally, no disialylated oligosaccharide chains occur on the Fc^{14} . The extent of oligosaccharide heterogeneity in a single crystal is identical to that found in pooled Fc fragments^{14,30}, consequently the X-ray data represent the composite structure. **c.** Fc fragment containing oligosaccharides devoid of galactose and sialic acid on each of the (α 1 $\rightarrow$ 6) arms. Since these residues in normal IgG are in contact with the surface of the protein (see b), their absence vacates oligosaccharide-binding sites in IgG from arthritic patients and could make the IgG 'sticky' by creating a lectin-like activity. It is not known to what extent the remaining sugar residues remain in contact with the peptide.

P-4 chromatograms of osteoarthritic (OA) patients are also characteristic of all such patients and are distinct from both the control and rheumatoid arthritic (RA) profiles (Fig. 3). Finally, the differences between the control and arthritic (both osteoarthritic and rheumatoid) P-4 profiles can be rationalized in terms of a population shift towards oligosaccharide structures of lower hydrodynamic volume. To establish the molecular basis

of this shift, the asialo oligosaccharides from each patient were analysed with respect to their relative levels of different core substitutions and the degree and nature of their outer-arm substitutions (see Figs 1, 4).

For the asparagine-linked oligosaccharides of IgG, core substitutions can be readily determined by digesting the pool of oligosaccharides with a mixture of *Streptococcus pneumoniae*

Table 1 Per cent oligosaccharide chains lacking galactose in serum IgG

Individual patients	Control		Osteoarthritis		Rheumatoid arthritis		Duration disease (yr)	Serology
Oxford series	Oxf1	26.7	Oxf7f*	38.9	Oxf13f	51.5	25	+
	Oxf2	23.9	Oxf8f*	36.0	Oxf14f	53.7	18	+
	Oxf3	26.2	Oxf9f	26.5	Oxf15f	41.9	2	+
	Oxf4	23.4	Oxf10m	35.2	Oxf16f	43.6	13	-
	Oxf5	16.5	Oxf11f	31.4	Oxf17f	54.9	3	+
	Oxf6	19.3	Oxf12f	33.0	Oxf18f	55.0	23	-
Tokyo series	Tok1f	31.4	Tok9f	32.2	Tok16m	44.3	4	+
	Tok2m	26.8	Tok10f	32.7	Tok17f	53.8	10	+
	Tok3m	38.9	Tok11f	32.7	Tok18f	52.4	16	+
	Tok4m	25.4	Tok12f*	47.4	Tok19f	48.4	13	+
	Tok5f	23.3	Tok13f	43.2	Tok20m	43.9	16	+
	Tok6f	19.4	Tok14f	33.4	Tok21f	47.9	18	+
	Tok7m	28.7	Tok15f	35.4				
	Tok8f	25.7						
Mixed series					O/T1m	74.5	18	+
					O/T2f	56.2	27	+
					O/T3m	55.6	>15	+
					O/T4f	51.8	15	+
					O/T5m	44.3	15	+
					O/T6m	38.7	>10	+
Mean $\pm$ s.d.	Oxf	22.7 $\pm$ 4.0		33.5 $\pm$ 4.3		50.1 $\pm$ 5.9		
	Tok	27.5 $\pm$ 5.8		36.7 $\pm$ 6.1		51.0 $\pm$ 9.1		
Pooled control serum	Bern-1†	24.0		—		—		

Non-galactosylated oligosaccharides were isolated as follows. Asialo oligosaccharides (1×10^7 c.p.m.) from the IgG of each individual were applied to an RCA-120 agarose column (Miles-Yeda, Lot no. AR26) of dimensions 0.6×20 cm. The column was developed in 5 mM sodium acetate (pH 5.6). Non-galactosylated structures eluted in the void while digalactosylated and monogalactosylated structures eluted later and at unique volumes (data not shown). The number of galactose residues in each peak was confirmed by following the change in hydrodynamic volume after digestion with jack bean β -N-acetyl hexosaminidase (14 μ M substrate, 150 U ml⁻¹ of enzyme in 0.1 M citrate phosphate, pH 4.5). For the Oxford series, the differences in the means were significant, with $P = 0.002$ (C vs OA) and $P < 0.001$ (C vs RA) and $P = 0.002$ (OA vs RA). For the Tokyo series (including data from Tok and O/T patients), the significance was $P = 0.006$ (C vs OA), $P < 0.001$ (C vs RA) and $P = 0.0008$ (OA vs RA). Any difference in the means between the two series was not considered statistically significant and no correlation was found between the extent of galactosylation and age and sex. The ratios of digalactosylated to monogalactosylated structures obtained as described in the text were as follows: C, 0.89 ± 0.09 ; OA, 0.75 ± 0.15 ; RA, 0.63 ± 0.12 . The statistical significance of these is $P = 0.07$ (C vs OA), $P = 0.0003$ (C vs RA), $P > 0.1$ (OA vs RA). Statistical analysis was performed using a non-parametric combined-order statistic test (Wilcoxon-Mann-Whitney test²⁵). Probabilities are quoted for a two-tailed test with a null hypothesis $H_0: \mu_1 = \mu_2$ tested against an alternative hypothesis $H_1: \mu_1 \neq \mu_2$. The mixed series refers to serum samples from British individuals, where the IgG was purified in England and sugar analysis performed in Tokyo. Blood samples in the Oxford series were obtained from patients at St John's Highfield Hospital, Droitwich, UK, and the Queen Elizabeth Medical Centre, Birmingham, UK. Patients Oxf13 to Oxf18, Tok16 to Tok21 and O/T1 to O/T6 had active disease and all fulfilled the American Rheumatism Association criteria for definite or classical rheumatoid arthritis²⁶. The Oxf patients range in age was 50–75 yr (mean, RA 64 ± 8 s.d.; OA, 68 ± 9 s.d.), and the Tok patients 41–75 yr (mean, RA 57 ± 14 s.d.; OA, 58 ± 9 s.d.). Control serum in the Oxf series was obtained from random blood bank donors. The Tok series control serum was from donors, age range 22–47 yr (mean 38 ± 9 s.d.). The analysis was performed double-blind, with clinical histories obtained after completion of oligosaccharide analysis.

* Patients Oxf7 and Oxf8 had long-standing osteoarthritis and very recently have been showing signs of an inflammatory component (6 months and 24 months respectively). The patient Tok 12 (aged 74 yr) was originally diagnosed as osteoarthritic but now also has Sjögren's syndrome.

† Bern-1 refers to IgG from the pooled serum of several thousand individuals and was given by Dr U. Nydegger of the Blood Transfusion Centre of the Swiss Red Cross, Bern, Switzerland.

β -galactosidase and β -N-acetylhexosaminidase¹³. The resulting digestion products are diagnostic for each of the four cores and differ sufficiently in hydrodynamic volume to be resolved on a P-4 column. The results (Fig. 4) indicate clearly that there is no systematic correlation between disease state and the incidence of any particular type of core structure (that is, no apparent changes in the levels of the enzymes GnT III (ref. 15) and $\alpha(1 \rightarrow 6)$ fucosyltransferase).

The outer-arm structures can be characterized with respect to the incidence, linkage and location of galactose, N-acetylglucosamine and N-acetylneuraminic acid. The asialo oligosaccharide mixtures were therefore subjected to *Ricinus communis* lectin-agarose affinity chromatography to separate galactosylated and non-galactosylated structures. Table 1 shows that in control IgG, ~75% of all oligosaccharide chains contain at least one galactose residue. In IgG isolated from rheumatoid and osteoarthritic patients, only ~50% ($P = 0.001$) and ~65% ($P = 0.002$), respectively, of the chains contained galactose. The ratio in each individual case of digalactosyl to monogalactosyl structures was determined either by chromatography of the galactosylated oligosaccharides on *R. communis* lectin-agarose, or enzymatically. In the latter, digestion with jack bean β -N-

acetylhexosaminidase, followed by P-4 gel permeation chromatography, results in the resolution of fragments diagnostic of the digalactosylated and monogalactosylated structures. The ratios of the digalactosyl to monogalactosyl structures obtained by both methods were consistent and indicated respective decreases of 30% ($P = 0.0003$) and 15% ($P = 0.07$) in the rheumatoid and osteoarthritic asialo oligosaccharide mixtures.

To determine whether the decreased number of chains containing galactose was secondary to a decrease in outer-arm $\beta(1 \rightarrow 2)$ linked N-acetylglucosamine residues, the asialo oligosaccharide structures from each individual were digested with jack bean β -galactosidase and the resulting degalactosylated oligosaccharides subjected to P-4 chromatography. In all three groups essentially identical profiles were found, implying no deficiency in outer-arm $\beta(1 \rightarrow 2)$ -linked N-acetylglucosamine residues (GlcNAc 5 and 5', Fig. 2). There can, therefore, be no difference between individuals, regardless of their disease state, with respect to the extent of N-acetylglucosaminylation (that is, no apparent change in the levels of enzymes GnT I and GnT II (ref. 15)).

As the negatively charged N-acetylneuraminic acid residues confer mobility in an electric field, an aliquot of the labelled

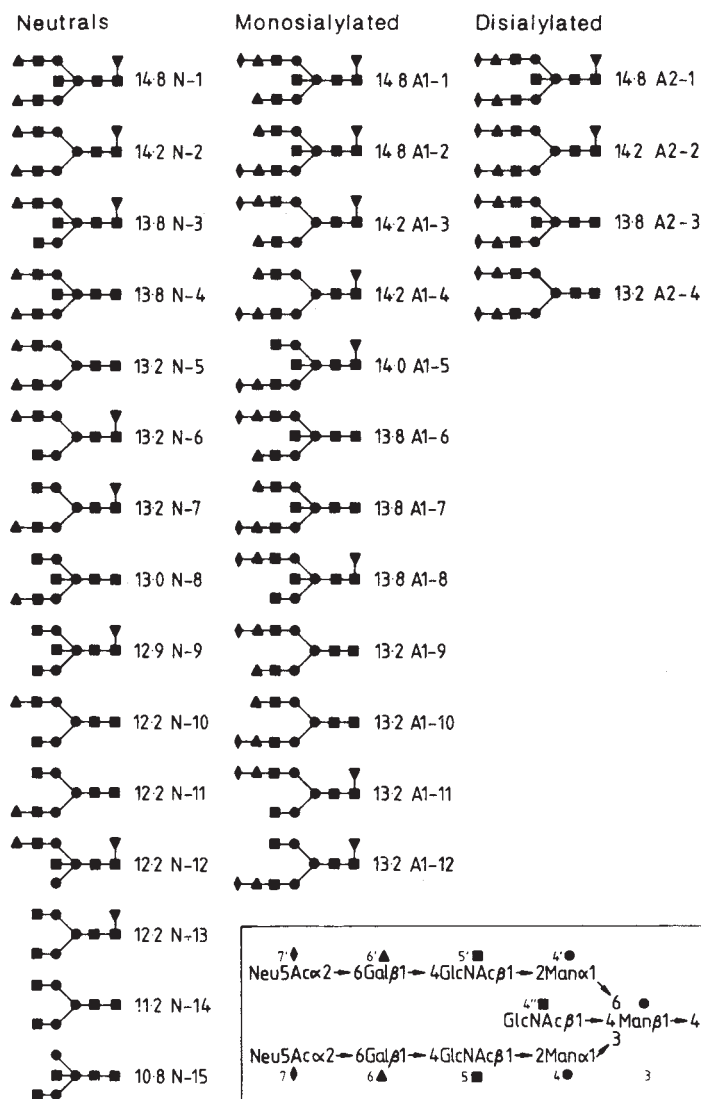


Fig. 2 Primary sequences of the *N*-linked oligosaccharides associated with human IgG. The hydrodynamic volume (as measured in glucose units) of each structure (or of its neutral derivative in the case of those sialylated) is indicated (refs 12, 13 and R.B.P. *et al.*, unpublished data), and was determined by comparison with α (1-6)-linked glucose oligomer standards.

oligosaccharide pool (before neuraminidase digestion) was subjected to high-voltage paper electrophoresis. The oligosaccharides were completely separated into neutral, monosialylated and disialylated structures. The occurrence of these structures is represented in Fig. 4 and the structures present in the three peaks are detailed in Fig. 2. Within each study the osteoarthritis and rheumatoid IgGs show a slight decrease in the number of chains terminating with one or two *N*-acetylneuraminic acid residues, consistent with their decrease in galactosylation.

The results show that osteoarthritis, and particularly rheumatoid arthritis, are associated with marked changes in the level of outer-arm galactosylation of the complex *N*-linked oligosaccharides of serum IgG ($P=0.002$ OA versus control (C), $P=0.001$ RA versus C) and these also distinguish the disease states ($P=0.002$ OA versus RA). Importantly, no novel oligosaccharides were found on IgG from either of the arthritides (data not shown).

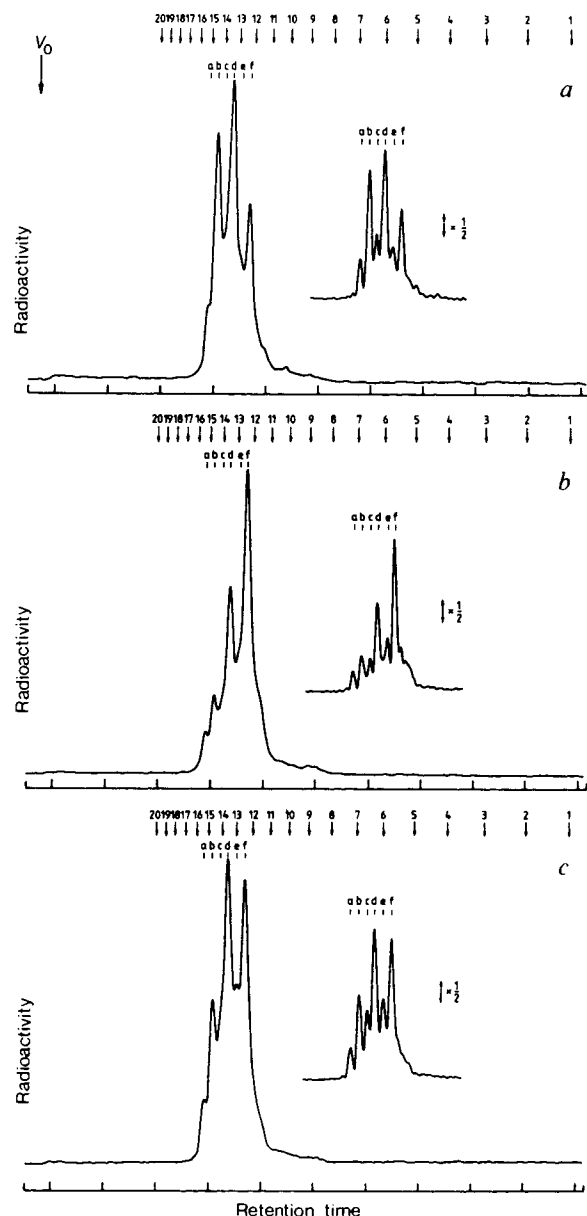
A possible consequence of this change in galactosylation is suggested by recent X-ray crystallographic studies on rabbit Fc, which have helped to define in molecular terms the manner in which the opposing oligosaccharides interact with each other and with the peptide¹⁶. In particular, the oligosaccharide chains appear to have different primary sequences (and hence many IgG molecules are structurally asymmetrical) and conformations (Fig. 1b). The α (1-6) arms interact with the protein surface, the major contacts being through their Neu5Ac α (2-6)Gal β (1-4)GlcNAc segment. The rigid¹⁷ α (1-3) arm of at least one of the paired oligosaccharides is always devoid of galactose,

thereby exposing its outer-arm β (1 $\rightarrow$ 2) linked *N*-acetylglucosamine residue, which then interacts directly with the Man β (1-4)GlcNAc segment of the opposing carbohydrate chain. The α (1-3) arm of the latter oligosaccharide extends outwards between the domains with no apparent steric constraints on its length (Fig. 1b).

Given this minimum restriction on pairing of oligosaccharides in Fc, a probability analysis reveals that the above changes in galactosylation would lead to a marked elevation in the incidence of Fc molecules that totally lack galactose (~300% in rheumatoid arthritis and ~90% in osteoarthritis, that is, from 10% in normal serum to 19% in osteoarthritis and 40% in rheumatoid arthritis). Thus, the hierarchical set of changes, beginning with an altered level of galactosylation and proceeding via a change in the relative populations of a constant set of oligosaccharide structures, leads through the phenomenon of pairing in the Fc, to dramatic changes in the incidence of individual Fc subpopulations. Such a change in IgG could result in the formation and/or persistence of immunoglobulin complexes by any one of several possible mechanisms that may or may not involve a true autoimmune response. First, oligosaccharides now terminating in *N*-acetylglucosamine could expose previously masked protein determinants or create new protein-oligosaccharide determinants that may be immunogenic. Second, an increase in the level of certain IgG subpopulations could lead to the exposure of certain Fc determinants at much higher concentrations than before. This could elicit a new immune response or raise a pre-existing low-level (subclinical)

Fig. 3 Representative Bio-Gel P-4 (~400 mesh) gel permeation chromatograms of the asialo oligosaccharides of total serum IgG: a, control; b, rheumatoid arthritis; c, osteoarthritis.

Methods. IgG from the serum of each of 42 individuals was isolated by precipitation at 4°C with ammonium sulphate (33%) and DE-52 ion-exchange chromatography (4°C) in 0.01 M NaPO₄ pH 7.2. In the Tokyo series, oligosaccharides of each purified IgG were released by hydrazinolysis and analysed by Bio-Gel P-4 column chromatography (after sialidase treatment) as reported previously¹³. In the Oxford series, each purified IgG (5 mg) was dialysed exhaustively against distilled water (4°C) and cryogenically dried over activated charcoal at -196°C (<10⁻⁶ bar). Protein samples were suspended in freshly distilled anhydrous hydrazine for 8 h at 100°C under an anhydrous argon atmosphere. The hydrazine was removed by repeated (5×) flash-evaporation from anhydrous toluene. The hydrazinolysates were N-acetylated by addition of an excess of acetic anhydride in saturated NaHCO₃ at 4°C. After passage through a column of Dowex AG 50×12 (H⁺ form), the samples were evaporated to dryness (27°C), re-dissolved in water and applied to Whatman 3 MM chromatography paper. Descending paper chromatography (27°C) was subsequently performed using n-butanol/ethanol/water (4:1:1 v/v) (solvent I). After 48 h the first 5 cm measured from the origin was eluted with H₂O. The oligosaccharides so isolated were flash-evaporated to dryness (27°C) and reduced with a fivefold excess of NaBH₄ (7.6 Ci mmol⁻¹, NEN) in 50 mM NaOH pH 12.0, 30°C for 4 h. An equivalent volume of 1 M NaBH₄ in 50 mM NaOH pH 12.0 was then added and incubation continued for 2 h. The mixture was then acidified (pH 5-6) with 1 M acetic acid and passed through a Dowex AG 50×12 (H⁺) column, evaporated to dryness (27°C) and flash-evaporated (27°C) from methanol (×5). The samples were then applied to Whatman 3 MM paper and subjected to descending paper chromatography for 2 days in solvent I. Radiochromatogram scanning was performed with a Berthold radiochromatogram scanner LB280. The radioactivity at the origin was subsequently eluted with water. An aliquot of reduced [³H]-labelled oligosaccharides so isolated was subjected to exhaustive neuraminidase digestion (*Arthrobacter ureafaciens*, Nakarai Chemical Co., Kyoto, Japan); radioactive sugar (1×10⁷ c.p.m.) in 50 µl of 0.1 M sodium acetate pH 5.0 containing 0.1 U of enzyme. Incubation was performed at 37°C for 18 h under a toluene atmosphere. The samples were then subjected to high-voltage paper electrophoresis at 80 V cm⁻¹ in pyridine/acetic acid/water (3:1:387 v/v, pH 5.4). All radioactivity remained at the origin (data not shown), confirming the complete cleavage of sialic acid. The samples were recovered from paper by elution with water, desalted using a tandem column of Dowex AG 50×12 (H⁺) and AG 3×4A (OH⁻) in water, evaporated to dryness, resuspended in 175 µl of a 20 mg ml⁻¹ partial dextran acid hydrolysate and applied to a Bio-Gel P-4 (~400 mesh) gel permeation chromatography column (1.5 cm×200 cm). The column was maintained at 55°C and water (200 µl min⁻¹) was used as the eluent. The eluent was monitored for radioactivity using a Berthold HPLC radioactivity monitor (model LB503) and for refractive index using a Perkin-Elmer model LC25 refractometer. Analog signals from the monitors were digitized using Nelson Analytical ADC interfaces. The digital values were collected and analysed using Hewlett Packard 9836C computers. The figure shows radioactivity (vertical axis) plotted against retention time after removal of stochastic noise using Fourier transform techniques. The numerical superscripts refer to the elution position of glucose oligomers in glucose units as detected simultaneously by the refractive index monitor (data not shown). V₀ is the void position. Sample elution positions (in glucose units) were calculated by cubic spline interpolation between the internal standard glucose oligomer positions. The insets are of resolution-enhanced profiles drawn with the retention time axis unchanged and the radioactivity axis reduced by 0.5. Resolution-enhancement was achieved by lineshape transformation in the Fourier domain.



response to a pathological one. Third, the oligosaccharide-binding sites in the C₂ domain, which are normally occupied by Neu5Ac(2-6)Gal segments of the (α1→6) arm, are now vacant and could effectively make the IgG 'sticky' by creating a lectin-like activity (see Fig. 1c). This could result in complex formation or auto-aggregation without an actual autoimmune response. Such a mechanism is implicit in some previous studies^{18,19} and in the occurrence of rheumatoid factor in apparently normal individuals. In particular, it would readily explain the weak affinity of binding between Fab and Fc in IgG complexes in rheumatoid serum^{8,20} and the unusual properties of many rheumatoid factors, for example, the monovalency of their interaction with IgG^{7,8} and their selective cross-reactivity with IgG from other species^{8,21,22}. None of the above mechanisms requires the oligosaccharide itself to be immunogenic and all are therefore consistent with the lack of any novel oligosaccharides on the IgG of rheumatoid arthritis patients.

Primary osteoarthritis is also associated with an altered glycosylation pattern of serum IgG, yet this disease is generally considered clinically distinct from rheumatoid arthritis and is not associated with pathological levels of immunoglobulin com-

plexes in the serum. Possibly the much smaller decrease in galactosylation of the IgG in osteoarthritis ensures that the level and size of any immunoglobulin complexes formed is too low to lead to a persistent inflammatory component, although sub-clinical levels of such complexes may exist.

As the glycosylation of IgG is altered, it is probable that the glycosylation of other glycoproteins of B lymphocytes (for example, cell-surface class I and class II major histocompatibility complex (MHC) antigens) and perhaps other cellular compartments is also affected. Susceptibility to many autoimmune diseases and some infections (for example, leprosy) correlates with a particular allelic form of one or more genes from one of the three HLA classes²³. In rheumatoid arthritis susceptibility is strongly correlated with the presence of the HLA-DRw4 allele^{23,24}. The correlation between decreased galactosylation and occurrence of rheumatoid arthritis raises the possibility that altered glycosylation of certain MHC-region products may be one of the means by which this genomic region can influence disease susceptibility.

We thank Professor R. R. Porter and Dr G. E. Paget for their interest and encouragement throughout this work. The Oxford

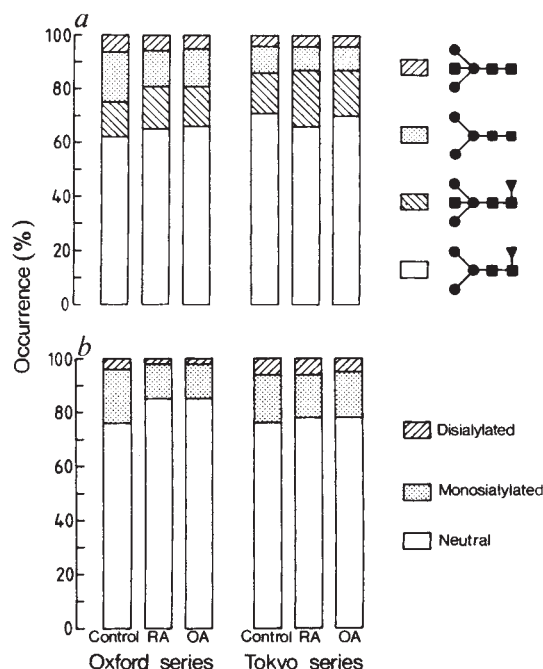


Fig. 4 *a*, Relative occurrence of the four core structures. An aliquot ($3-5 \times 10^5$ c.p.m.) of the unfractionated asialo oligosaccharides was digested with a mixture of *S. pneumoniae* β -galactosidase (2 mU) and β -N-acetylhexosaminidase (4 mU) in 25 μ l of 0.1 M citrate phosphate pH 6.0. The digestion was performed at 37 °C for 18 h under a toluene atmosphere and terminated by heating to 100 °C for 2 min. After desalting [Dowex AG 50 $\times$ 12 (H^+), AG3 $\times$ 4A (OH^-)], the digestion products were separated on a Bio-Gel P-4 (~400 mesh) column. Enzymes were purified by a modification of the method of Glasgow *et al.*³¹ *S. pneumoniae* β -galactosidase removes all galactose residues from the asialo oligosaccharides of IgG, irrespective of their core structures (data not shown). *S. pneumoniae* β -N-acetylhexosaminidase digests the resulting oligosaccharides in a manner dependent on the presence of the (bisecting) GlcNAc β 1 $\rightarrow$ 4 residue³². Specifically, only one N-acetylglucosamine residue (GlcNAc 5 in Fig. 2) is released from the agalactosyl structure GlcNAc β 1 $\rightarrow$ 2 Man α 1 $\rightarrow$ 6 (GlcNAc β 1 $\rightarrow$ 2 Man α 1 $\rightarrow$ 3) (GlcNAc β 1 $\rightarrow$ 4) Man β 1 $\rightarrow$ 4 GlcNAc β 1 $\rightarrow$ 4 ($\pm$ Fuc α 1 $\rightarrow$ 6) GlcNAc, which is converted to GlcNAc β 1 $\rightarrow$ 2 Man α 1 $\rightarrow$ 6 (Man α 1 $\rightarrow$ 3) (GlcNAc β 1 $\rightarrow$ 4) Man β 1 $\rightarrow$ 4 GlcNAc β 1 $\rightarrow$ 4 ($\pm$ Fuc α 1 $\rightarrow$ 6) GlcNAc. However, two N-acetylglucosamine residues are released from GlcNAc β 1 $\rightarrow$ 2 Man α 1 $\rightarrow$ 6 (GlcNAc β 1 $\rightarrow$ 2 Man α 1 $\rightarrow$ 3) Man β 1 $\rightarrow$ 4 GlcNAc β 1 $\rightarrow$ 4 ($\pm$ Fuc α 1 $\rightarrow$ 6) GlcNAc, which is converted to Man α 1 $\rightarrow$ 6 (Man α 1 $\rightarrow$ 3) Man β 1 $\rightarrow$ 4 GlcNAc β 1 $\rightarrow$ 4 ($\pm$ Fuc α 1 $\rightarrow$ 6) GlcNAc. The four cores have the following structures: $\blacksquare$, (+B-F), Man α 1 $\rightarrow$ 6 (Man α 1 $\rightarrow$ 3) (GlcNAc β 1 $\rightarrow$ 4)Man β 1 $\rightarrow$ 4GlcNAc β 1 $\rightarrow$ 4GlcNAc; $\blacksquare$, (-B-F), Man α 1 $\rightarrow$ 6(Man α 1 $\rightarrow$ 3) Man β 1 $\rightarrow$ 4GlcNAc β 1 $\rightarrow$ 4GlcNAc; $\blacksquare$, (+B+F), Man α 1 $\rightarrow$ 6(Man α 1 $\rightarrow$ 3) (GlcNAc β 1 $\rightarrow$ 4) Man β 1 $\rightarrow$ 4GlcNAc β 1 $\rightarrow$ 4 (Fuc α 1 $\rightarrow$ 6)GlcNAc; $\square$, (-B-F), Man α 1 $\rightarrow$ 6(Man α 1 $\rightarrow$ 3) Man β 1 $\rightarrow$ 4GlcNAc β 1 $\rightarrow$ 4 (Fuc α 1 $\rightarrow$ 6)GlcNAc. The percentage of each core as determined in the Oxford series was: +B-F: C, 6.6 $\pm$ 2.5; OA, 5.1 $\pm$ 2.2; RA, 6.0 $\pm$ 2.2; -B-F: C, 18.5 $\pm$ 6.5; OA, 13.0 $\pm$ 4.3; RA, 14.3 $\pm$ 3.1; +B+F: C, 13.2 $\pm$ 3.1; OA, 15.1 $\pm$ 3.9; RA, 15.8 $\pm$ 5.0; -B+F: C, 61.8 $\pm$ 10.6; OA, 66.8 $\pm$ 7.2; RA, 63.9 $\pm$ 6.0. The values determined in the Tokyo series were: +B-F: C, 3.8 $\pm$ 2.7; OA, 3.7 $\pm$ 0.8; RA, 4.6 $\pm$ 1.4; -B-F: C, 8.8 $\pm$ 2.9; OA, 9.7 $\pm$ 2.4; RA, 8.8 $\pm$ 2.1; +B+F: C, 14.6 $\pm$ 1.9, OA, 16.6 $\pm$ 4.0; RA, 20.9 $\pm$ 5.7; -B+F: C, 72.8 $\pm$ 3.3; OA, 70.1 $\pm$ 5.1, RA, 65.7 $\pm$ 6.7. There is no statistical significance ($P > 0.05$) in the different incidence of any core in the three groups. *b*, Relative occurrence of oligosaccharide chains containing terminal α (2-6)-linked N-acetylneuraminic acid. An aliquot (5×10^5 c.p.m.) of 3H -labelled oligosaccharides was subjected to high-voltage paper electrophoresis (80 V cm $^{-1}$) in pyridine/acetic acid/water (3:1:387 v/v) pH 5.4 (Camag HVE cell 61000). After scanning with a Berthold LB280 radiochromatogram scanner, the regions N (neutral), A-1 (first acidic) and A-2 (second acidic) were eluted and the radioactivity in each determined, allowing the ratios in each sample to be calculated. To confirm that all structures present in the A-1 and A-2 positions contained only one and two sialic acid residues respectively, an aliquot of the eluted A-1 and A-2 oligosaccharide fractions was applied to a QAE-A25 Sephadex column (6 mm $\times$ 10 cm). Samples were applied in 2 mM ammonium acetate (pH 5.3) and the column developed with a 2-350 mM linear gradient of ammonium acetate. Both the A-1 and A-2 fractions from the paper electrophoresis gave single elution peaks corresponding to the standard elution positions of monosialylated and disialylated oligosaccharides, respectively. The percentage of each as determined in the Oxford series was: N: C, 75.7 $\pm$ 5.0; OA, 84.7 $\pm$ 1.8; RA, 85.2 $\pm$ 1.1; A-1: C, 20.5 $\pm$ 5.0; OA, 13.7 $\pm$ 1.5; RA, 12.6 $\pm$ 1.1; A-2: C, 3.8 $\pm$ 1.6; OA, 1.7 $\pm$ 0.7; RA, 2.2 $\pm$ 0.6. The values determined in the Tokyo series were: N: C, 76.1 $\pm$ 1.1; OA, 78.6 $\pm$ 2.1; RA, 78.4 $\pm$ 4.4; A-1: C, 17.8 $\pm$ 1.7; OA, 16.5 $\pm$ 1.8; RA, 15.2 $\pm$ 3.2; A-2: C, 6.1 $\pm$ 1.5, OA, 5.0 $\pm$ 0.9; RA, 6.5 $\pm$ 1.5.

Oligosaccharide Group is supported by the Monsanto Company. R.A.D., T.W.R. and B.J.S. are members of the Oxford Enzyme Group. The Tokyo Group is supported by research grants for Scientific Research, from the Ministry of Education, Science and Culture of Japan, and by a research grant from the Yamanouchi Foundation for Research on Metabolic Disorders.

Received 23 January; accepted 30 May 1985.

- Kunkel, H. G. & Tan, E. M. *Adv. Immunol.* **4**, 351-395 (1964).
- Pope, R. M., Mannik, M., Gilliland, B. C. & Teller, D. C. *Arthritis Rheum.* **18**, 97-106 (1975).
- Stuart, J. M., Townes, A. S. & Kang, A. H. *Rev. Immunol.* **2**, 199-218 (1984).
- Shakib, F. & Stanworth, D. R. *Ann. rheum. Dis.* **37**, 12-17 (1978).
- Solinger, A. M., Bhatnagar, R. & Stobo, J. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3877-3881 (1981).
- Klareskog, L., Forsum, U., Scheynius, A., Kabelitz, D. & Wigzell, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3632-3636 (1982).
- Winchester, R. J., Kunkel, H. G. & Agnello, V. J. *exp. Med.* **134**, 2865 (1971).
- Nardella, F. A., Teller, D. C. & Mannik, M. J. *exp. Med.* **154**, 112-125 (1981).
- Henney, C. S. & Stanworth, D. R. *Nature* **201**, 511-512 (1964).
- Mullinax, F., Hymes, A. J. & Mullinax, G. L. *Arthritis Rheum.* **19**, 813 (1976).
- Takasaki, S., Mizuuchi, T. & Kobata, A. *Meth. Enzym.* **83**, 263-268 (1982).
- Yamashita, K., Mizuuchi, T. & Kobata, A. *Meth. Enzym.* **83**, 105-126 (1982).
- Mizuuchi, T., Taniguchi, T., Shimizu, A. & Kobata, A. *J. Immunol.* **129**, 2016-2020 (1982).
- Rademacher, T. W. & Dwek, R. A. *Prog. Immunol.* **5**, 95-112 (1983).
- Yamashita, K., Tachibana, Y., Ohkura, T. & Kobata, A. *J. biol. Chem.* **260**, 3963-3969 (1985).
- Sutton, B. J. & Phillips, D. C. *Biochem. Soc. Trans.* **11**, 130-132 (1982).
- Homans, S. W., Dwek, R. A., Fernandes, D. L. & Rademacher, T. W. *FEBS Lett.* **164**, 231-235 (1983).
- Heimer, R., Fenton, M. R. & Abruzzo, L. J. *Immunochimistry* **8**, 603-611 (1971).
- Hymes, A. J., Mullinax, G. L. & Mullinax, F. J. *biol. Chem.* **254**, 3148-3151 (1979).
- Normansell, D. E. *Immunochimistry* **8**, 593-602 (1971).
- Normansell, D. E. & Stanworth, D. R. *Immunology* **15**, 549-560 (1968).
- Normansell, D. E. & Stanworth, D. R. *Immunology* **10**, 527-533 (1966).
- Panayi, G. S. in *Immunogenetics* (eds Panayi, G. S. & David, C. S.) Ch. 6 (Butterworths, London, 1984).
- McMichael, A. J., Sasazuki, T., McDevitt, H. O. & Payne, R. O. *Arthritis Rheum.* **20**, 1037-1042 (1977).
- Lloyd, E. *Handbook of Application Mathematics* Vol. 6, Pt B (Wiley, New York, 1984).
- Primer on the Rheumatic Diseases* 7th edn, Appendix 1 (ed. Rodman, G. P.) 137-140 (Arthritis Foundation, New York, 1973).
- Sox, H. C. Jr. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **66**, 975-982 (1970).
- Spiegelberg, H. L., Abel, A. C., Fishkin, B. G. & Grey, H. M. *Biochemistry* **9**, 4217-4223 (1970).
- Montreuil, J. *Biochem. Soc. Trans.* **11**, 134-136 (1982).
- Rademacher, T. W. *et al. Biochem. Soc. Trans.* **11**, 132-134 (1982).
- Glasgow, L. R., Paulson, J. C. & Hill, R. L. *J. biol. Chem.* **252**, 8615-8623 (1977).
- Yamashita, K., Ohkura, T., Yoshima, H. & Kobata, A. *Biochem. biophys. Res. Commun.* **100**, 226-232 (1981).

SV40 enhancer and large-T antigen are instrumental in development of choroid plexus tumours in transgenic mice

Richard D. Palmiter*, Howard Y. Chen†, Albee Messing‡ & Ralph L. Brinster†

* Howard Hughes Medical Institute, Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

† Laboratory of Reproductive Physiology, School of Veterinary Medicine and ‡ Division of Neuropathology, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

We have shown recently that choroid plexus tumours frequently develop in transgenic mice which have developed from fertilized eggs injected with DNA molecules containing both simian virus 40 (SV40) early-region genes and metallothionein (MT) fusion genes¹, and several lines of mice have now been established in which all of the offspring that inherit the foreign DNA succumb to these tumours at 3-5 months of age (ref. 1 and our unpublished data). Several other tissues, notably thymus and kidney, occasionally also show pathological changes. SV40 large-T antigen protein and messenger RNA are always present in affected tissues at much greater concentrations than in unaffected tissues, suggesting that SV40 early-region genes are preferentially activated in choroid plexus, thymus and kidney and that this activation frequently leads to tumorigenesis in the choroid plexus¹. To determine which regions of the original constructs are important for this tumorigenesis, we have now tested several derivatives and report here that the large-T antigen is sufficient, that the MT fusion gene is dispensable and

that the SV40 enhancer (72-base-pair repeat region) has an important role in directing tumours to the choroid plexus. Deletion of the SV40 enhancer region alone commonly leads to peripheral neuropathy, as well as liver and pancreatic tumours, which are the subject of the accompanying paper. Evidence is presented that these pathologies may result from an enhancing effect of the *MT* sequences on large-T antigen genes, made possible by removal of the otherwise dominant SV40 enhancer.

Figure 1 depicts the various DNA constructs that were compared in the present study. The top line shows one of the original constructs, 143 (or SV-MGH), which nearly always results in choroid plexus tumours. It consists of a *KpnI/BamHI* fragment which includes the SV40 enhancer, 21-bp repeats, the 'TATA-box' and early-region genes (coding for small-T and large-T antigens) as well as a *KpnI/EcoRI* fragment containing the MGH gene, which has ~650 bp of *MT* gene 5'-flanking sequence and promoter fused to the human growth hormone (*hGH*) structural gene. The SV40 early-region genes and MGH genes are transcribed in opposite directions. The SV40 early region gives rise to two mRNAs by differential splicing; one codes for large-T antigen while the other codes for small-T antigen³. Constructs containing both of the early-region genes (188 and 189) were compared with deletion mutants that can code only for intact large-T antigen (pSV11) or for intact small-T antigen and a truncated large-T antigen (pSV8). We also compared the effect of deleting the enhancer (*KpnI/SphI*) and enhancer plus 21-bp repeat regions *KpnI/NcoI* from these early-region genes using constructs in which the MGH fusion gene was either present (144 and 202) or absent (203 and 204).

In each case, several hundred copies of linearized DNA fragments containing the various gene constructs were excised from plasmids and microinjected into the male pronuclei of fertilized one-cell F₂ hybrid eggs (obtaining by mating C57 × SJL hybrid adults); the eggs were then transferred to the oviducts of pseudopregnant foster mothers for further development. Offspring carrying the microinjected fusion genes were identified by dot hybridization of nucleic acids isolated from tail biopsies. The mice were monitored carefully and killed for histological examination and analysis of SV40 gene expression when any abnormalities were detected. The results are summarized in Fig. 1 and Table 1.

In our original study, 10 mice with the SV-MGH construction were studied; 8 of them died prematurely. Choroid plexus papillomas were positively identified in three of these mice; the other five were not autopsied but probably also died of brain tumours¹. An additional eight transgenic mice with the same 143 construct were produced for this study. Seven of these mice developed choroid plexus papillomas; three of them also showed signs of thymic hyperplasia, one had nephritis, one had polycystic kidneys and one had an hepatocellular carcinoma (Table 1). Thus, the spectrum of abnormalities in these mice resembles that reported previously and confirmed that choroid plexus tumours are the primary cause of death when foreign DNA containing both SV40 early genes and a MGH fusion are present in all cells of a mouse.

To ascertain whether the MGH gene sequences are involved in directing tumours to the choroid plexus, these sequences were deleted in constructs 188 and 189. Construct 188 is wild-type SV40, whereas 189 has a 20-bp insertion near the 3' end of the large-T antigen gene¹. Because these two constructs are functionally equivalent, they will be considered together here. Eight of nine mice that carried only the SV40 early-region genes died of choroid plexus papillomas and the spectrum of other abnormalities resembled those obtained with the 143 construct (Table 1). Polycystic kidneys were particularly frequent. Thus, the MGH fusion genes are neither essential nor influential when present in the context of construct 143.

To distinguish between choroid plexus pathology resulting from large-T antigen, small-T antigen or both, we tested two internal deletions in the SV40 early region: one (pSV11) encodes only an intact large-T antigen and the other (pSV8) encodes an intact small-T antigen and a truncated large-T antigen⁴. Four of seven mice with the pSV11 construct died with choroid plexus papillomas; thymic hyperplasia and kidney lesions were also apparent in some of these mice. One mouse died with kidney glomerular lesions but also had a small choroid plexus tumour; offspring of this mouse had both choroid plexus tumours and glomerular pathology. One of nine mice with the pSV8 construct died of hydrocephalus; this is relatively common in this strain of mice and is likely to be unassociated with the pSV8 insert. A second mouse carrying pSV8 sequences died at 1 yr of age as a result of nonspecific skin lesions that frequently occur in

Table 1 Pathology found in mice containing various SV40 gene constructs

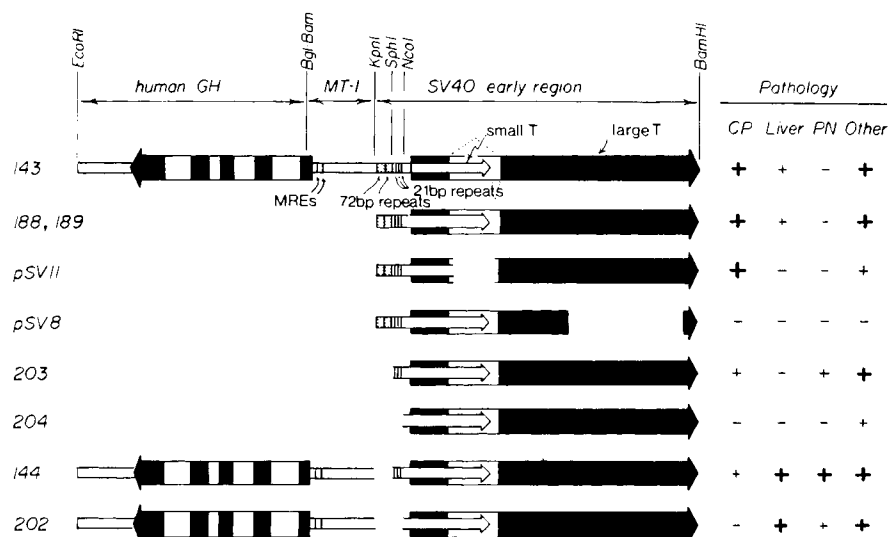
Gene construct	No. of mice produced	No. of mice autopsied	Location of pathology					
			CP	Thy	Kid	PN	Liv	Panc
143 SV-MGH	8	8	7	3	2	0	1	0
188 SV (wt)	5	5	5	4	4	0	0	0
189 SV (ts)*	4	4	3	1	1	0	1	0
pSV11 (T)	7	6	5	1	1	0	0	0
pSV8 (t)	9	2	0	0	0	0	0	0
203 SVΔe	12	11	1	1	0	2†	0	3
204 SVΔe + 21	7	3	0	0	0	0	0	1
144 SVΔe-MGH	16	16	1	0	0	13	11	8
202 SVΔe + 21-MGH	6	6	0	0	1	3†	5	3

Overt pathology: CP, choroid plexus papillomas; Thy, thymic hyperplasia; Kid, kidney lesions; PN; peripheral neuropathy; Liv, liver carcinomas; Panc, pancreatic islet cell adenomas. The primary cause of death with the different constructs was probably as follows. Construct 143, seven mice died with CP tumours, one had kidney failure; construct 188, all five mice died with CP tumours; construct 189, three mice died with CP tumours, one with hydrocephalus; pSV11, four mice died with CP tumours, one died of kidney failure (this mouse also had a small CP tumour and its offspring died of CP tumours), one mouse died unexpectedly at about 1 month with no apparent symptoms and one mouse is still alive after 9 months; pSV8, one mouse died with hydrocephalus and another died with no apparent symptoms, seven mice are still alive after 7 months; construct 203, three mice died with pancreatic tumours, one with a CP tumour, one with an undifferentiated brain tumour, one with weak hind legs, one with no apparent symptoms, one was killed at 2 months as a control for PN, three were killed while apparently healthy but beyond the age when most mice develop tumours and one is still alive after 12 months; construct 204, one mouse died with a pancreatic tumour, two with no apparent symptoms and four are still alive after 12 months; construct 144, 16 mice died of generalized paralysis from PN, liver or pancreatic tumours (one of these also had a small CP tumour); construct 202, five mice died of liver or pancreatic tumours, one mouse died unexpectedly at about 1 month with no apparent symptoms.

* We originally thought that this construct was a temperature-sensitive derivative of SV40 (ts58), but subsequent analysis showed that it was not temperature-sensitive.

† These peripheral neuropathies were milder than most of those seen in mice carrying construct 144 (see accompanying paper² for details).

Fig. 1 Diagrams of the various SV40 constructs and a summary of their associated pathology. The construction of SV-MGH has been described elsewhere¹; the vector sequences are not shown; the locations of restriction sites used in the construction of SV-MGH and various deletions are indicated. In the *hGH* region, the exons are solid; the metal regulatory elements (MREs) of the mouse *MT-I* promoter region are indicated; the 72-bp repeats which include the SV40 enhancers are stippled; the 21-bp repeats are shown as bars; the SV40 early region codes for two mRNAs, small-T (the open inner arrow) and large-T (the solid outer arrow), which are generated by differential splicing (dotted lines). Constructs 143, 188, 189, 202, 203 and 204 were prepared by deleting the indicated sequences; pSV8 and pSV11 were constructed by Colby and Shenk⁴. In each case, the plasmids were linearized within the pBR322 vector before injection into fertilized mouse eggs. A summary of the major pathology is given in Table legend.



control animals of that age. No other abnormalities were noted. Thus, small-T antigen plus the truncated large-T antigen produced by pSV8 appear to be insufficient to promote tumorigenesis in this model system. In contrast, large-T antigen alone (pSV11) is sufficient to induce the characteristic spectrum of abnormalities associated with the SV40 early region, in agreement with results obtained by injecting mutant SV40 viruses into hamsters⁵.

At least one copy of the 72-bp repeats (the 'enhancer region') of SV40 is essential for viral replication⁶ and sequences within this region have been shown to enhance transcription of various heterologous genes in a position- and orientation-independent manner⁷⁻⁹. Between the 72-bp repeats and the TATA-box lie three 21-bp repeats that bind a specific transcription factor¹⁰; these are implicated in efficient transcription of both early and late genes.¹¹ Thus, we decided to delete these sequences as well. In construct 203, only the 72-bp enhancer region has been deleted from the SV40 early genes. Six of twelve mice with this construct died with notable pathology; only one had a choroid plexus papilloma; three had pancreatic tumours, one had an enlarged thymus, one had an undifferentiated brain tumour (not of the choroid plexus) and two had a mild form of peripheral neuropathy. Thus, the frequency of choroid plexus tumours is significantly reduced when the enhancer is deleted, and the overall frequency of pathology also appears to be reduced. When both the enhancer and the 21-bp repeat regions were deleted (construct 204), only one of seven mice died with notable pathology. In this case there were tumours in the pancreas and myometrium. Thus, it appears that deletion of both of these control regions severely reduces the ability of the SV40 early-region genes to produce tumours.

In another set of experiments, we deleted the SV40 control regions but left the *MGH* gene intact. The results with the 144 construct (which deletes only the enhancer) revealed a new spectrum of abnormalities. Thirteen of 16 mice that retained construct 144 showed signs of peripheral neuropathy. Typically, these mice showed hind limb paralysis, or weakness, sometimes as early as 3 weeks of age. In addition, 11 of these mice had hepatocellular carcinomas, 8 had pancreatic islet cell adenomas and 1 had a choroid plexus papilloma (see accompanying paper²). Deletion of both the enhancer and 21-bp repeat regions (construct 202) led to a similar distribution of pathologies, including peripheral neuropathy, and liver and pancreatic tumours (Table 1). The results of these internal deletions are strikingly different from those obtained with the comparable constructs described above which lack the *MGH* gene. In the absence of the SV40 control elements, the site of pathology appears to be redirected to peripheral nerves and organs such as liver and pancreas. Furthermore, SV40 genes that were almost

completely ineffective at producing tumours (construct 204) appear to be effective when the *MGH* gene is present (construct 202). Thus, we tentatively conclude that sequences within the *MGH* region have an important role in directing the site of pathogenesis when the SV40 control elements are deleted.

To ascertain whether large-T antigen was synthesized even in those mice carrying constructs lacking the normal SV40 control regions, cell lines were isolated from hepatocellular tumours of mice with the 143, 144 and 202 constructs. When these cells were fixed and incubated with antisera directed against SV40 large-T antigen followed by a second fluorescein-labelled antibody, the nuclei fluoresced brightly, indicating that large-T antigen was being synthesized and transported to the nucleus in all of these cell lines (Fig. 2a), even 144 and 202 which lack important SV40 promoter/enhancer elements.

The SV40 mRNA transcripts were mapped by primer extension to establish whether the normal start sites were being used. When a 22-nucleotide (nt) primer was hybridized to RNA from tumour cells derived from mice with the 143 or 144 constructs and elongated with reverse transcriptase, two products of ~100 bp were produced, similar to the pattern observed with SV40-infected cells¹⁰. However, in two cell lines from mice with the 202 construct, these extension products were a minor component and a heterogeneous mixture of transcripts that originate within the *MT* flanking region was seen (Fig. 2b). To quantitate mRNA production, replicate cultures were incubated for 10 h with or without the addition of 10 μ M cadmium to the medium;

Table 2 Effects of heavy metals on SV40 and *MGH* gene expression

Cell line	DNA construct	Metal	SV40 early mRNA (molecules per cell)	<i>MGH</i> mRNA
345-3	143	None	270	400
		10 μ M Cd	275	925
517-5	144	None	55	540
		10 μ M Cd	150	1,470
70-4	202	None	120	4,000
		10 μ M Cd	350	10,740

Cell lines were isolated from mice with the SV-MGH constructs indicated in Fig. 1. The cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. For the experiments described here, the medium was changed and 10 μ M CdSO₄ was added; 10 h later the cells were collected and total nucleic acids isolated. The amount of *MGH* and SV40 mRNA was determined by solution hybridization using ³²P-labelled cDNA and M13 standards essentially as described by Durnam and Palmiter²⁰.

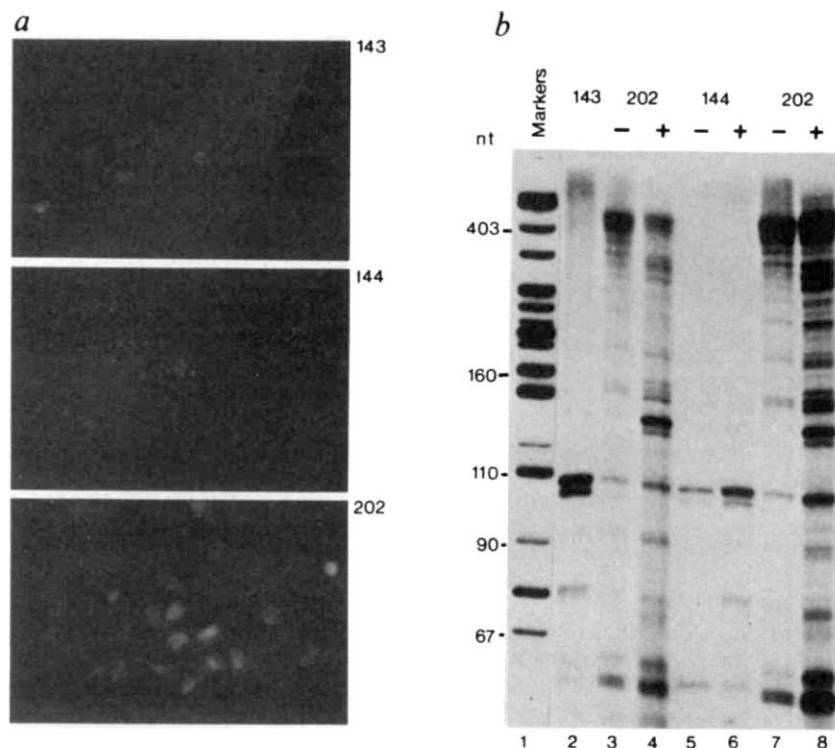


Fig. 2 Expression of SV40 large-T antigen and its mRNA in cell lines derived from tumour cells of transgenic mice. **a**, Immunofluorescence of SV40 large-T antigen in hepatocellular carcinoma cell lines derived from mice with the 143, 144 or 202 gene constructs shown in Fig. 1. **b**, Primer extension analysis of SV40 RNA from various cell lines. The 22-nucleotide primer (5' GCTGCAAGATTCCTCTCTGTT 3') was labelled with T4 kinase and [γ - 32 P]ATP, annealed with 20 μg total RNA and extended with avian myeloblastosis virus reverse transcriptase. The extension products were separated on a denaturing 8% acrylamide gel. Markers are end-labelled *Hpa*II digests of pBR322. Lane 1, markers; lane 2, hepatocellular carcinoma cells (construct 143; mouse 53-4); lanes 3, 4, pancreatic cells (construct 202, mouse 76-6); lanes 5, 6, hepatocellular carcinoma cells (construct 144; mouse 517-5); lanes 7, 8, hepatocellular carcinoma cells (construct 202, mouse 70-4). Cells were grown without (-) or with (+) 80 μM Zn and 8 μM Cd for 14 h before preparing RNA for analysis.

Methods. **a**, Cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. Cells were fixed on the plastic dishes with 3.6% formaldehyde/1% methanol, rinsed well, then incubated with anti-large-T antigen, followed by fluorescein isothiocyanate-labelled goat anti-hamster immunoglobulin. The fluorescence was visualized with a Leitz microscope.

total nucleic acids were then isolated and assayed for the amount of *MGH* and SV40 mRNA by solution hybridization. Both SV40 and *MGH* mRNAs were induced severalfold when cadmium was added to the medium of cells from tumours carrying construct 144 or 202 (Table 2). However, SV40 mRNA was generally uninducible by metals in tumour cells from mice with the 143 construct, in which the SV40 enhancers are intact. In cell lines from mice with the 144 construct, transcripts with the normal start site were induced, whereas in cell lines from mice with the 202 construct, transcripts either shorter or longer than normal were also induced by heavy metals (Fig. 2b). The fact that SV40 genes respond to heavy-metal induction after the SV40 enhancers have been deleted suggests that the *MT* promoter sequence, which carries several metal regulatory elements¹², is acting as a metal-dependent enhancer to the SV40 transcription unit which lies ~600 bp away. This may explain the preponderance of liver tumours with constructs 144 and 202, because liver is one of the major sites of MT synthesis¹³. Peripheral neuropathy may also be a consequence of MT influence, although MT expression has never been examined in peripheral nerves.

In conclusion, those constructs in which the SV40 enhancer region and large-T antigen gene were intact have a strong propensity to direct tumours to the choroid plexus of mice. The large-T antigen appears to be sufficient to promote all of the pathological consequences of the intact early region; thus, the role of small-T antigen remains uncertain even in this model system, where the genes and their products have access to every possible cell type. Removal of the SV40 enhancer leads to various other pathologies; the location of these pathologies may be attributed to the remaining control regions, the 21-bp repeats for example. The simplest interpretation of the enhancer deletion mutations, in which the *MGH* gene was retained, is that the early-region genes then come under the influence of the *MT* regulatory sequences. With this view, the SV40 early-region genes are sensitive to external enhancers if their own enhancer is deleted. It is not clear whether the SV40 enhancer is intrinsically dominant over other enhancer-like sequences or whether its proximity to the early-region genes overrides the *MT* sequence effects. The *MT* genes are 'housekeeping' genes which

are expressed in many different cells¹³; thus, their enhancer-like functions are also likely to be expressed in many different cells. Several cell-specific enhancers have recently been identified¹⁴⁻¹⁶ and some of these have already been shown to direct tumours to unique cell types¹⁷⁻¹⁹. Thus, microinjection of various oncogenes associated with selective cellular enhancers into fertilized eggs should allow insight into the role of different oncogenes in the same cellular environment or the same oncogene in different cellular environments.

We thank Arnold Levine and Terry van Dyke for supplying the pSV11 and pSV8 DNAs that were originally constructed by W. Colby and T. Shenk. We also thank Myrna Trumbauer, Paul Fallon and Mary Yagle for technical assistance, J. Peschon for help with primer extension analysis, G. S. McKnight for immunofluorescence, and our colleagues for their helpful suggestions during this study. The work was supported by NIH research grants CA 38635, HD 09172 and NS 05572. A.M. is a recipient of a postdoctoral fellowship from the Smith Kline Beckman Pathology Education Fund.

Received 4 January; accepted 4 June 1984.

1. Brinster, R. L. *et al.* *Cell* **37**, 367-379 (1984).
2. Messing, A., Chen, H. Y., Palmiter, R. D. & Brinster, R. L. *Nature* **316**, 461-463 (1985).
3. Toozé, J. *DNA Tumor Viruses. The Molecular Biology of Tumor Viruses* 2nd edn, Pt 2 (Cold Spring Harbor Laboratory, New York, 1980).
4. Colby, W. W. & Shenk, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5189-5193 (1982).
5. Lewis, A. M. Jr & Martin, R. G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4299-4302 (1979).
6. Gruss, P., Dhar, R. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 943-947 (1981).
7. Banerji, J., Rosconi, S. & Schaffner, W. *Cell* **27**, 299-308 (1981).
8. Gluzman, Y. & Shenk, T. (eds) *Enhancers and Eukaryotic Gene Expression* (Cold Spring Harbor Laboratory, New York, 1983).
9. Wasylyk, B., Wasylyk, C. & Chambon, P. *Nucleic Acids Res.* **12**, 5589-5608 (1984).
10. Dynan, W. & Tjian, R. *Cell* **32**, 669-680 (1983).
11. Vigneron, M. *et al.* *EMBO J.* **3**, 2373-2382 (1984).
12. Stuart, G. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 7318-7322 (1984).
13. Kagi, H. R. & Nordberg, M. (eds) *Metallothionein* (Birkhauser, Basel, 1979).
14. Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729-739 (1983).
15. Walker, M. D. *et al.* *Nature* **306**, 557-561 (1984).
16. Ornitz, D. M. *et al.* *Nature* **313**, 600-603 (1985).
17. Stewart, T. A., Pattengale, P. K. & Leder, P. *Cell* **38**, 627-637 (1984).
18. Hanahan, D. *Nature* **315**, 115-122 (1985).
19. Ornitz, D. M. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **50** (in the press).
20. Duram, D. M. & Palmiter, R. D. *Analyt. Biochem.* **131**, 385-393 (1983).

Peripheral neuropathies, hepatocellular carcinomas and islet cell adenomas in transgenic mice

Albee Messing*§, Howard Y. Chen†, Richard D. Palmiter‡ & Ralph L. Brinster†

* Division of Neuropathology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA

† Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

‡ Howard Hughes Medical Institute, Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

§ Present address: School of Veterinary Medicine, University of Wisconsin, Madison, Wisconsin 53706, USA

The ability to introduce foreign DNA into the genome of mice offers unique opportunities to produce new models of disease processes. Recent experiments have shown that integration and expression of simian virus 40 (SV40) T antigen genes¹ and the murine mammary tumour virus (MMTV)-*myc* genes² in transgenic mice can lead to the development of neoplasia in a remarkably tissue-specific manner. In the case of SV40-bearing mice, tumours consistently develop in the choroid plexus³. In the accompanying paper, we show that the 72-base pair (bp) enhancer in the SV40 genome is instrumental in directing tumorigenesis to the choroid plexus³. However, when the enhancer is deleted from a construction also containing the metallothionein-human growth hormone fusion gene (SVΔe-MGH), an entirely new pattern of pathology results. The present report focuses on transgenic mice carrying this construct; they develop demyelinating peripheral neuropathies, hepatocellular carcinomas and islet cell adenomas.

Fertilized mouse eggs were injected with several hundred copies of the SVΔe-MGH fusion gene (referred to as gene construct 144 in the accompanying paper³); the eggs were then transferred to foster mothers for further development. A total of 16 offspring carried the fusion gene. Of these mice, 13 suffered from varying degrees of weakness or paralysis, which became apparent as early as 2–3 weeks of age. Mice were killed after they began to show signs of illness, and tissues were examined for gross and histological abnormalities. The results are summarized in Table 1.

To identify the cause of the paralysis, several regions along the neuromuscular axis (brain, spinal cord, peripheral nerves, muscle) were examined histologically. Consistent lesions were found in the spinal roots and peripheral nerves. In paraffin-embedded sections from severely affected mice, virtually all

Table 2 Serum glucose and insulin levels in SVΔe-MGH mice with islet cell adenomas of the pancreas

Gene construction	Mouse no. and sex	Age* (days)	Serum glucose (mg per 100 ml)	Serum insulin† (ng ml ⁻¹)
SVΔe-MGH	103-2♂	110	8	4.9
SVΔe-MGH	137-6♀	78	2	20.6
SVΔe-MGH	99-6-6♀	59	3	10.0
SVΔe-MGH	136-1♂	72	6	215
SVΔe-MGH	136-1-1♂	75	2	240
SVΔe-MGH	136-1-9♀	69	44	34.8
SVΔe-MGH	136-1-1♂	42	63	1.2
SVΔe-MGH	136-1-9♀	42	69	0.5
MGH‡	4♂, 4♀	54-102	169 ± 12§	4.3 ± 0.8
Controls	1♂, 8♀	59-164	203 ± 20	0.6 ± 0.1

* Age when blood samples were taken for glucose and insulin determinations.

† Serum insulin levels were determined by radioimmunoassay against rat insulin standards.

‡ Four of these mice contained the metallothionein-rat growth hormone gene construct¹⁹, and four contained the metallothionein-human growth hormone construct²⁰. Note that the serum insulin levels in the MGH mice are higher than in the controls; nevertheless, these mice are able to maintain normal serum glucose levels. The values for the MGH mice are more appropriate as a baseline from which to judge the levels seen in the SVΔe-MGH mice with islet cell adenomas.

§ Mean ± 1 s.e.

spinal roots, peripheral nerves and intramuscular nerve branches showed increased numbers of cell nuclei. To characterize the nerve lesions in more detail, samples from the mid-thigh sciatic nerves of 14 mice were fixed in glutaraldehyde and embedded in epoxy for light and electron microscopy. For comparison, Fig. 1a shows a normal sciatic nerve from an SV-MGH mouse (referred to as gene construct 143 in ref. 3), which contains the same gene construct as SVΔe-MGH but with the 72-bp enhancer intact. Nerves from severely affected SVΔe-MGH mice showed increased numbers of large and irregular Schwann cell nuclei, including some mitotic figures, and a marked paucity of myelin (Fig. 1b). In longitudinal sections segmental demyelination was seen. There was no evidence of active axonal degeneration or of damage to the neuronal cell bodies of the ventral spinal motoneurons and peripheral ganglia. The sciatic nerves of mildly affected mice had more normally myelinated large-diameter fibres, but still showed scattered axons with abnormally thin myelin sheaths (Fig. 1c). These features are most consistent with a primary demyelinating neuropathy⁴.

At the electron microscopic level, the Schwann cells possessed basement membranes and demonstrated appropriate initial

Table 1 Incidence of pathology in SVΔe-MGH transgenic mice

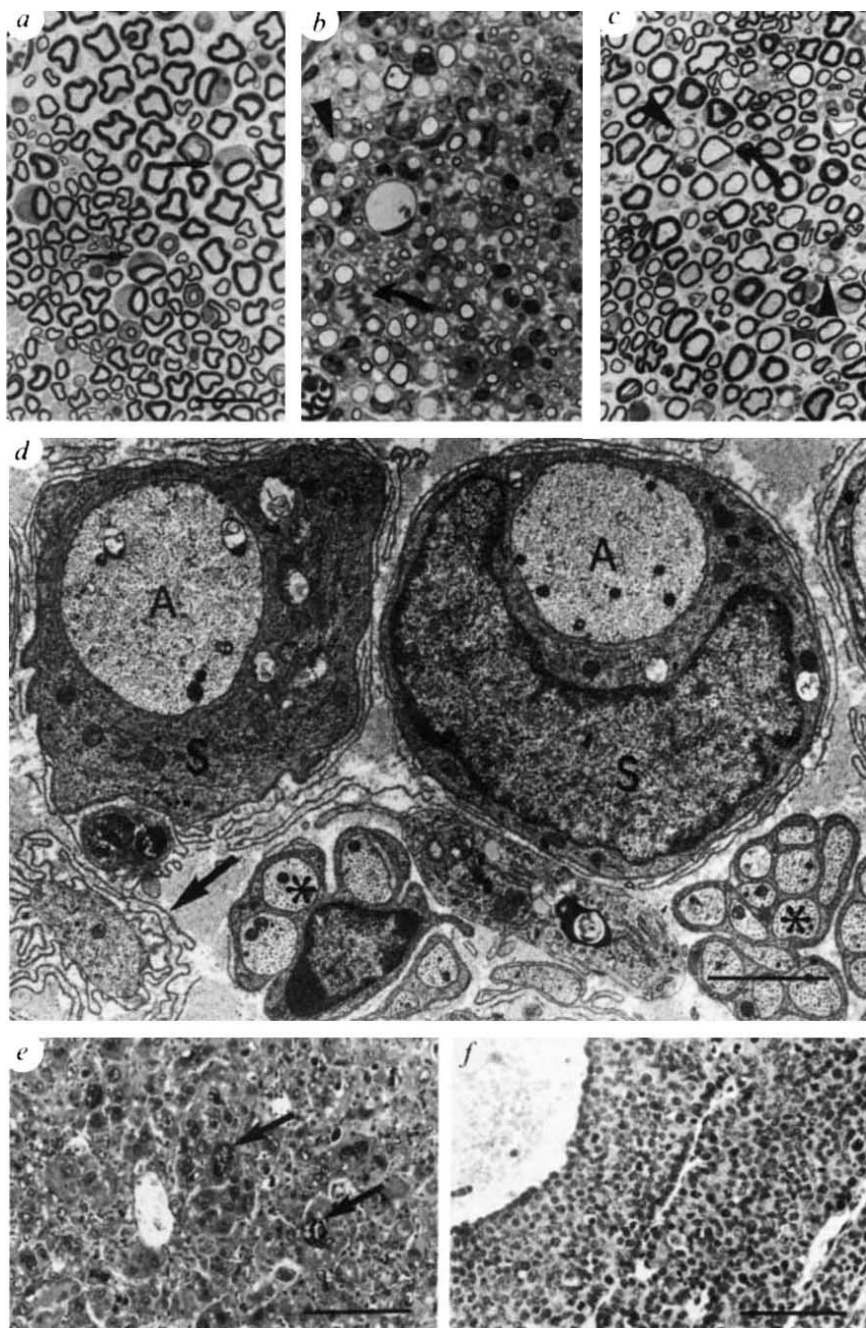
Expt	Transgenic mice	SV40 DNA* copies per cell	Mean lifespan (days)	No. of mice with pathology			
				Peripheral neuropathy†	Hepatic carcinoma	Islet cell adenoma	Choroid plexus papilloma‡
I	2	1.2-1.7	115	1	2	0	0
II	3	9.3-112	42	2	1	0	0
III	8	0.9-41	85	7	6	5	1
IV	3	1.0-10.2	70	3	2	3	0
Totals	16			13	11	8	1

*SV40 gene copy number per cell was determined on samples of total cellular nucleic acid from tail biopsies as described previously¹⁸.

† Number of mice with histologically confirmed lesions in peripheral nerves. Samples of mid-thigh sciatic nerve were processed as described in Fig. 1a legend and were scored at the light microscopic level by two independent observers as being either normal or having a mild, moderate or severe peripheral neuropathy, according to the degree of myelin loss. Examples of typical severely or mildly affected nerves are shown in Fig. 1b and c, respectively. Of the 13 mice with lesions in their peripheral nerves, 2 were judged severe, 4 were judged moderate, and 7 were judged mild. One mouse lived to 7 months without showing signs of weakness and had normal peripheral nerves, although it did have a hepatocellular carcinoma. Two mice with hindleg weakness died unexpectedly and the nerve were too autolysed for analysis.

‡ Miscellaneous lesions seen at low frequency include: three adrenal tumours, one pituitary carcinoma, one intramuscular haemangiosarcoma, four dysplasias of the transitional epithelium in the renal pelvis and ureters, and four instances of diffuse nuclear atypia in adipocytes of brown fat. These lesions have not been investigated in detail.

Fig. 1 Pathology in SVΔe-MGH transgenic mice. **a**, Light micrograph of transverse section of normal sciatic nerve from 40-day-old SV-MGH transgenic mouse with the gene construct containing the 72-bp enhancer. After sodium pentobarbital anaesthesia of the mouse, its sciatic nerve was exposed and fixed *in situ* with 3.6% glutaraldehyde in 0.1 M sodium phosphate buffer, pH 7.4 for 10 min and then removed, placed in fresh fixative, and stored at 4 °C for 2 days, washed twice with buffer, post-fixed in 1% osmium tetroxide in buffer for 2 h at room temperature, rewashed, dehydrated through alcohols and propylene oxide and embedded in epoxy resin. A 1-μm section was stained with toluidine blue. Many myelinated axons and several Schwann cell nuclei are seen in this section (two are indicated by arrows). The myelin is the concentric dark ring around each axon. Scale bar, 25 μm. **b**, Light micrograph of transverse section of sciatic nerve from 40-day-old SVΔe-MGH mouse that was one of the most severely affected. Note increased numbers of Schwann cell nuclei (one is indicated by straight arrow) and paucity of myelin around the large-diameter axons (one is indicated by the arrowhead). Many nuclei have more condensed chromatin and are larger and more irregular than in control nerve. A mitotic Schwann cell is indicated by the curved arrow. A few very thinly myelinated fibres are also present. Inflammatory cells are rare and only small amounts of myelin debris are seen. Longitudinal sections show numerous examples of segmental demyelination. Scale as in **a**. **c**, Light micrograph of transverse section of sciatic nerve from 73-day-old SVΔe-MGH mouse, which was one of the most mildly affected. The nerve was processed as in **a**. There are substantially more myelinated axons than in the nerve from the severely affected mouse shown in **b**, but there are still scattered fibres with abnormally thin myelin sheaths (one indicated by curved arrow) or with no myelin at all (two indicated by arrowheads). In longitudinal sections some myelin ovoids were seen. Scale as in **a**. **d**, Electron micrograph (JEOL 100-X) of 1,000-Å sections (LKB ultramicrotome), stained with uranyl acetate and lead citrate, from nerve shown in **b**. Two large-diameter axons (A) are shown enveloped by their own Schwann cells (S), but devoid of any loose or compact myelin. Axons of this diameter would normally be ensheathed by ~50 lamellae of compact myelin²¹. The Schwann cells possess a basement membrane; numerous strands of redundant basement membrane in the interstitium suggest Schwann cell degeneration (arrow). The axons appear ultrastructurally normal, except for a slight reduction in their diameter. The small-diameter axons and their associated Schwann cells (which typically have multiple axons associating with a single non-myelinating Schwann cell) appear normal (asterisks). Scale bar, 2 μm. **e**, Light micrograph of a section of liver from SVΔe-MGH mouse with hepatocellular carcinoma, paraffin-sectioned and stained with haematoxylin/eosin. The hepatocytes are variable in size and shape with extremely pleomorphic nuclei (arrows). Some livers, including this one, are grossly relatively normal but exhibit a diffuse type of atypia; others are grossly enlarged and infiltrated by multiple nodular masses, which contain hepatocytes growing in trabeculae or sheets with many mitoses. Both types of changes were classified as hepatocellular carcinomas. Scale bar, 200 μm. **f**, Light micrograph of islet cell adenoma from SVΔe-MGH mouse, paraffin-sectioned and stained with haematoxylin/eosin. These tumours vary from 2 mm to over 1 cm in diameter and are often multiple. The masses contain relatively uniform cells with round nuclei and granular eosinophilic cytoplasm, growing in small clusters on a thin capillary stroma. Many cystic spaces contain pooled blood or proteinaceous debris (upper left corner). The tumours are quite variable with respect to mitotic activity. Scale bar, 100 μm.



envelopment of the large-diameter axons. However, the Schwann cells seemed unable to form or maintain compact myelin (Fig. 1d). The small-diameter axons and their normally non-myelinating Schwann cells seemed completely normal. Preliminary indications are that myelination in the central nervous system by oligodendrocytes was also normal.

Lesions also developed in other tissues of several SVΔe-MGH mice, particularly the liver and pancreas (Table 1). In 11 mice the liver was diffusely infiltrated by white nodules of varying size, which on histological examination were classified as

hepatocellular carcinomas (Fig. 1e). In eight mice, masses were observed in the pancreas, all of which were classified as islet cell adenomas (Fig. 1f).

We further characterized the metabolic effects of the islet cell adenomas, with particular attention to the possible secretion of insulin, by measuring serum glucose and insulin levels. Blood samples were taken from six SVΔe-MGH mice with islet cell adenomas, just before the animals were killed, and serum glucose levels were found to be profoundly low in all but one mouse (Table 2). Serum insulin levels were determined on the

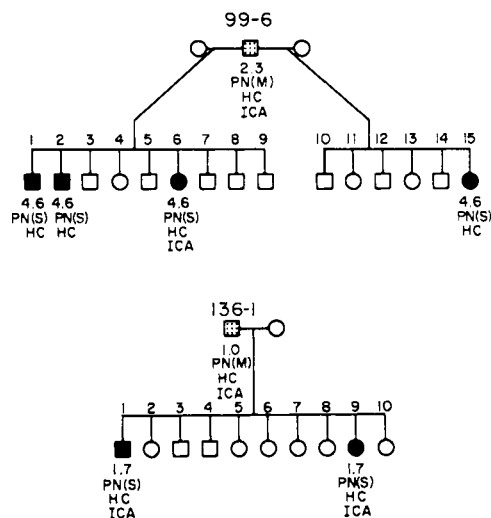


Fig. 2 Inheritance of neuropathy and other lesions in progeny from SVΔe-MGH transgenic mice. Pedigree of three litters obtained from two SVΔe-MGH males. Squares represent males and circles represent females. Closed symbols indicate mice that were positive for SV40 DNA (as determined by dot hybridization). The mouse number is shown above each symbol. Beneath each DNA-positive mouse is shown the DNA copy number per cell (determined by quantitative dot hybridization) and the types of lesions found at autopsy: PN, peripheral neuropathy (S, severe; M, mild); HC, hepatocellular carcinoma; ICA, islet cell adenoma of pancreas. Note that the gene transmission ratio to the offspring was <50% from both males, and for each parent all offspring showed a consistent increase of DNA copy number over that in the parent. For these reasons we believe the two parent males were actually mosaics, as indicated by the stippling within their symbols.

same samples by radioimmunoassay and were found to be elevated in five of six mice (Table 2). Preliminary immunohistochemical studies show that glucagon- and somatostatin-producing cells did not contribute to the tumour masses.

We considered whether the peripheral neuropathy could arise as an indirect effect of the islet cell adenomas. In two offspring that had islet cell adenomas at autopsy, serum glucose and insulin levels were determined at the first signs of paralysis (mice 136-1-1 and 136-1-9); at this time insulin levels were normal and glucose levels only moderately depressed (see Table 2). The lack of concordance between the onset of the peripheral neuropathy and the changes in serum glucose and insulin levels suggests that the peripheral neuropathy is unlikely to be secondary to insulin-induced hypoglycaemia. Further, the demyelinating nature of the neuropathy and the lack of apparent damage to neuronal cell bodies is not consistent with hypoglycaemic damage^{5,6}.

In contrast to SV-MGH mice³, choroid plexus papillomas were rare (1 out of 16) in SVΔe-MGH mice (Table 1). This low incidence partially reflects the shorter lifespan in SVΔe-MGH mice, since choroid plexus papillomas were not found in several SV-MGH mice killed at less than 2 months of age (our unpublished observations). However, if we consider only the SVΔe-MGH mice that lived longer than 2 months, a reduced incidence of choroid plexus papillomas was still observed (1 in 12 mice). For comparison, seven out of seven of the original SV-MGH mice that were allowed to reach this age, and six out of seven offspring from two of the originals, developed choroid plexus papillomas.

To determine whether the SVΔe-MGH construct was integrated into the host chromosomes in a manner that would allow transmission of the phenotype, three litters were produced from mice with the mild form of the neuropathy (weakness, with nerves as in Fig. 1c). These were the only mice that could live to reproductive age and mate normally. Six offspring inherited the SVΔe-MGH construct (Fig. 2), and all showed the severe form of peripheral neuropathy (paralysis, with nerves

as in Fig. 1b). The offspring were killed when they were 32–75 days old; all had hepatocellular carcinomas, three had islet cell adenomas with elevated serum insulin levels, and none had choroid plexus papillomas (Fig. 2, Table 2).

We were surprised to find that mice with the mild form of the neuropathy would produce offspring with the severe form. However, the gene was inherited by fewer than half the progeny in each pedigree, and the gene copy number per cell was higher in the offspring than in the parents (Fig. 2), thus establishing that both parents were in fact mosaics⁷. The mosaicism probably affects all somatic tissues (T. Wilkie, R.D.P. and R.L.B., unpublished observations); thus, the milder peripheral neuropathy in the parents may reflect the fact that only some (about half) of their Schwann cells carry the SVΔe-MGH construct.

We have shown that removal of the 72-bp enhancer element from the SV-MGH construct markedly reduces the incidence of choroid plexus papillomas in transgenic mice, but instead produces mice with high incidences of demyelinating peripheral neuropathy, hepatocellular carcinoma and islet cell adenomas. This change in the pattern of pathology implies an underlying change in the pattern of T antigen gene expression. While we might have expected liver pathology to occur in the event of the metallothionein promoter assuming control over expression of the SV40 early genes³, we do not understand the basis for the consistent lesions seen in pancreas and peripheral nerves.

The demyelinating polyneuropathy that we have observed in SVΔe-MGH transgenic mice strongly suggests a defect in the Schwann cell-axonal relationship, which normally leads to the formation and maintenance of myelin around large-diameter fibres⁸ and superficially resembles the peripheral nerve lesions in the *trembler* mouse mutant^{9,10} and Dejerine-Sottas neuropathy in man^{11,12}; however, the 'onion bulbs' characteristic of these hypertrophic neuropathies have not been seen in our mice. An SV40-related papovavirus causes the human central nervous system demyelinating disease progressive multifocal leukoencephalopathy^{13–15}; implication of SV40-related papovaviruses in other human neurological diseases remains controversial^{16,17}. Future studies will address the questions of which cell type in the SVΔe-MGH transgenic mice expresses the SV40 gene products and how this expression disrupts the process of myelination in the murine peripheral nervous system.

We thank Drs W. F. Hickey and J. T. Sladky for reviewing the histopathology, Myrna Trumbauer and Mary Yagle for technical assistance, and our colleagues for their helpful suggestions during the preparation of this manuscript. Radioimmunoassays for insulin were performed by the Diabetes Research Center at the University of Pennsylvania. This work was supported in part by NIH grants HD-09172, CA-38635 and NS-05572. A.M. was the recipient of a postdoctoral fellowship from the Smith-Kline Beckman Pathology Education Fund.

Received 4 January; accepted 4 June 1985.

- Brinster, R. L. *et al.* *Cell* **37**, 367–379 (1984).
- Stewart, T. A., Pattengale, P. K. & Leder, P. *Cell* **38**, 627–637 (1984).
- Palmiter, R. D., Chen, H. Y., Messing, A. & Brinster, R. L. *Nature* **This issue**.
- Dyck, P. J., Karnes, J., Lais, A., Lofgren, E. P. & Stevens, J. C. in *Peripheral Neuropathy* (eds Dyck, P. J., Thomas, P. K., Lambert, E. H. & Bunge, R.) 760–870 (Saunders, Philadelphia, 1984).
- Jaspan, J. B., Wollman, R. L., Bernstein, L. & Rubenstein, A. H. *Medicine* **61**, 33–44 (1982).
- Sidenius, P. & Jakobsen, J. *Diabetes* **32**, 383–386 (1983).
- Palmiter, R. D., Wilkie, T. M., Chen, H. Y. & Brinster, R. L. *Cell* **36**, 869–877 (1984).
- Bray, G. M., Rasminsky, M. & Aguayo, A. J. *Rev. Neurosci.* **4**, 127–162 (1981).
- Ayers, M. M. & Anderson, R. M. *Acta neuropath.* **25**, 54–70 (1973).
- Low, P. A. & McLeod, J. G. *J. neurol. Sci.* **26**, 565–574 (1975).
- Dyck, P. J. & Gomez, M. R. *Mayo Clin. Proc.* **43**, 280–296 (1968).
- Dyck, P. J. in *Peripheral Neuropathy* (eds Dyck, P. J., Thomas, P. K., Lambert, E. H. & Bunge, R.) 1630–1634 (Saunders, Philadelphia, 1984).
- Zu Rhein, G. M. & Chou, S. *Science* **148**, 1477–1479 (1965).
- Padgett, B. L., Zu Rhein, G. M., Walker, D. L., Eckroade, R. J. & Dessel, B. H. *Lancet* **i**, 1257–1260 (1971).
- Weiner, L. P. *et al.* *New Engl. J. Med.* **286**, 385–390 (1972).
- Johnson, R. T. in *Polyomaviruses and Human Neurological Diseases* (eds Sever, J. L. & Madden, D. L.) 183–190 (Liss, New York, 1983).
- Farwell, J. R., Dohrmann, G. J. & Flannery, J. T. *J. Neurosurg.* **61**, 657–664 (1984).
- Palmiter, R. D., Chen, H. Y. & Brinster, R. L. *Cell* **29**, 701–710 (1982).
- Palmiter, R. D. *et al.* *Nature* **300**, 611–615 (1982).
- Palmiter, R. D., Norstedt, G., Gelinas, R. E., Hammer, R. E. & Brinster, R. L. *Science* **222**, 809–814 (1983).
- Friede, R. L. & Samorajski, T. *J. comp. Neurol.* **130**, 223–232 (1967).

Two tandemly organized human genes encoding the T-cell γ constant-region sequences show multiple rearrangement in different T-cell types

Marie-Paule Lefranc* & T. H. Rabbitts

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The recent detailed analysis of genes that undergo rearrangement in T cells has shown that the T-cell receptor genes encoding α - and β -chains are involved in specific alterations in T-cell DNA analogous to the immunoglobulin genes¹⁻⁹. A third type of gene, designated γ , has been isolated from mouse cytotoxic T lymphocytes¹⁰, and evidence suggests that the mouse displays very limited diversity in this gene system¹¹, having only three variable-region (V) genes and three constant-region (C) genes¹². The function of the so-called T-cell γ gene is unknown. We have isolated genomic genes encoding the human homologue of the mouse T-cell γ gene; as there is no evidence that this T-cell rearranging gene is anything to do with the T3 molecule, we have designated the human T-cell rearranging gene as TRG γ (ref. 13), to avoid confusion with the T3 γ -chain, and have shown that the gene locus maps to chromosome 7 in humans¹³. We now report that human DNA contains two tandemly arranged TRG γ constant-region genes about 16 kilobases apart. These two genes show multiple rearrangement patterns in a variety of T cells, including helper and cytotoxic/suppressor type, as well as in all forms of T-cell leukaemia. Our results indicate variability of this T-cell gene system in man compared with the analogous system in mouse.

The human counterparts of the mouse γ gene were isolated from a genomic library cloned in λ phage by hybridization with a γ complementary DNA clone (designated T γ 5) prepared from the mouse T-cell line EL4 (ref. 14). Initially one clone was isolated from a phage library of the Burkitt's lymphoma cell line Raji¹⁵. The restriction map of this clone (λ R γ) is shown in Fig. 1; the clone starts within the γ coding region and covers 12 kilobases (kb) upstream of it. A *Hind*III restriction fragment of 2.1 kb immediately upstream of the coding region was subcloned in an M13 vector (clone designated M13H60) and used as a probe in filter hybridizations with genomic DNA. A set of different non-T-cell genomic DNA samples was digested with *Bam*HI and hybridized with the M13H60 probe (Fig. 2a). Each sample showed the presence of two hybridizing fragments of 15 and 12.5 kb and there was therefore no evidence for restriction fragment length polymorphism connected with the *Bam*HI fragments hybridizing with this clone. As there is no site for the

*Bam*HI restriction enzyme within the genomic *Hind*III fragment used for the probe, there must be at least two non-allelic DNA segments encoding TRG γ -like sequences. This conclusion was confirmed by isolating a second set of phage clones from a phage library prepared with genomic DNA of a B-cell prolymphocytic leukaemia, D-PLL. These clones encompass two TRG γ constant regions separated by 16 kb. A restriction map is shown in Fig. 1a. The region immediately upstream from the two TRG γ coding regions and these overlapping clones show similarities in restriction sites and include *Hind*III fragments like that cloned in M13H60 (these homologous regions are indicated in the diagram within Fig. 1).

The two gene coding regions, designated TRG γ 1 and γ 2, were assigned within the restriction map of the clones by hybridization to the T γ 5 cDNA clone (Fig. 1) and each gene was found to possess an internal *Bam*HI site. The restriction mapping also shows that the genomic *Bam*HI fragments of 15 and 12.5 kb corresponded to TRG γ 1 and γ 2, respectively. A fourth clone was isolated from the phage library using the M13H60 probe and the restriction map of this clone (λ D7) is also shown in Fig. 1. This clone is identical with λ D11 except for the absence of the *Hind*III site corresponding to the 5' end of the M13H60 cloned fragment (this site is indicated in parentheses in Fig. 1a). The allelic nature of this gene was demonstrated by a filter hybridization experiment using various *Hind*III-digested DNA samples hybridized with the M13H60 probe (Fig. 2b). The panel of DNA samples used showed either the 2.1-kb *Hind*III fragment, which corresponds to the 2.1-kb fragment cloned in M13H60, or this fragment together with a 4.5-kb fragment corresponding to the non-allelic polymorphic fragment carried by clone λ D7. These results indicate that allelic chromosomes exist in humans, one which carries two 2.1-kb *Hind*III restriction fragments (allelic chromosome 1, Fig. 1) and one which carries a 2.1-kb fragment plus a 4.5-kb fragment (allelic chromosome 2, Fig. 1). This conclusion is supported by the inheritance pattern observed in the family shown in Fig. 2 (Fig. 2b, samples A, B and C). The mother (sample B) is homozygous for allelic chromosome 1 and the father (sample A) has at least one allelic chromosome 2. The offspring (sample C) has inherited a homozygous pattern for chromosome 1; the father (sample A) must therefore be heterozygous for allelic chromosomes 1 and 2.

Part of the TRG γ coding region was mapped to the 2-kb *Hind*III/*Bam*HI fragment (Fig. 1) by hybridization with the T γ 5 cDNA probe (M.-P.L., unpublished data). Partial nucleotide sequencing data were obtained for the TRG γ 1 gene from two of the λ phage clones (λ R γ isolated from Raji DNA and λ D19 isolated from D-PLL DNA) by subcloning this fragment in M13 vectors followed by sequencing with dideoxy chain termination procedures¹⁶. These sequences were aligned using the mouse γ sequences¹⁰ for positioning within the predicted coding sequence and show the start of the derived constant-region amino-acid sequence, with an RNA splice site, at nucleotide 22 (Fig. 3).

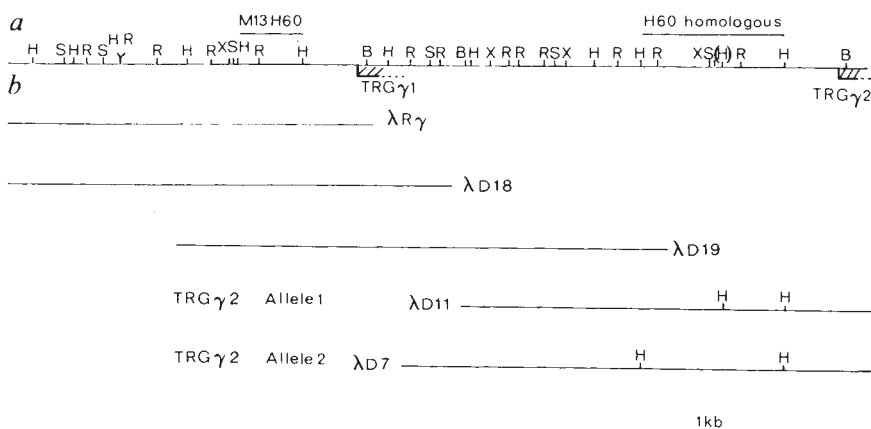


Fig. 1 Restriction map of human TRG γ genes. *a*, The upper line represents the partial map of two contiguous TRG γ genes. The nucleotide sequence obtained for the region adjacent to the *Bam*HI site within the coding region of TRG γ and the M13H60 probe are indicated. The homologous region upstream of the TRG γ 2 gene is marked and the polymorphic *Hind*III restriction site is indicated by parentheses upstream of TRG γ 2. *b*, The various λ phage clones are illustrated to show the areas of the γ locus which they cover. Note λ D11 carries the TRG γ 2 gene from allelic chromosome 1 and λ D7 carries the TRG γ 2 gene from allelic chromosome 2; the relevant polymorphic *Hind*III sites are indicated in these phages. H, *Hind*III; B, *Bam*HI; R, *Eco*RI; S, *Sac*I; X, *Xho*I.

* Permanent address: Laboratoire d'Immunogénétique, Université des Sciences et Techniques du Languedoc, 34060 Montpellier Cedex, France.

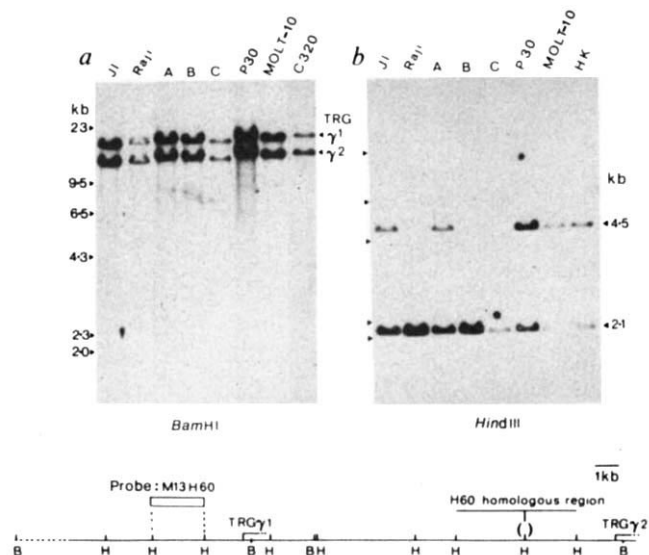


Fig. 2 Restriction fragment length polymorphism within human TRG γ genes. Genomic DNA from the cell lines indicated, from peripheral blood of a family (individuals designated A (father), B (mother) and C (son)) and an unrelated person (HK) were digested to completion with *Bam*HI (a) or *Hind*III (b), fractionated on 0.8% agarose gel and transferred to cellulose nitrate²². The immobilized DNA was hybridized as described elsewhere¹⁹ to nick-translated M13H60 (the *Hind*III fragment was subcloned in M13 mp8; ref. 23) and the hybridized genes visualized by autoradiography after washing excess probe. Sizes were estimated by co-electrophoresis of λ phage DNA digested with *Hind*III. The lower panel is a diagram of the human TRG γ constant-region locus indicating the cloned probe (M13H60) from the TRG γ 1 region and the homologous (polymorphic) region from the TRG γ 2 region. (The polymorphic *Hind*III site is shown in parentheses.) The area at the 5' end of each TRG γ gene is indicated over the *Bam*HI site. The cell lines used are JI and Raji, which are Burkitt's lymphoma B-cell lines^{24,25}; P30/Ohkubo and MOLT-10 are presumptive pre-T-cell lines²⁰ but we have shown that neither DNA has T-cell receptor β -chain genes rearranged and P30 has two immunoglobulin heavy-chain *J* alleles rearranged¹⁹. Colo320 (C320) is derived from a colon carcinoma¹⁷.

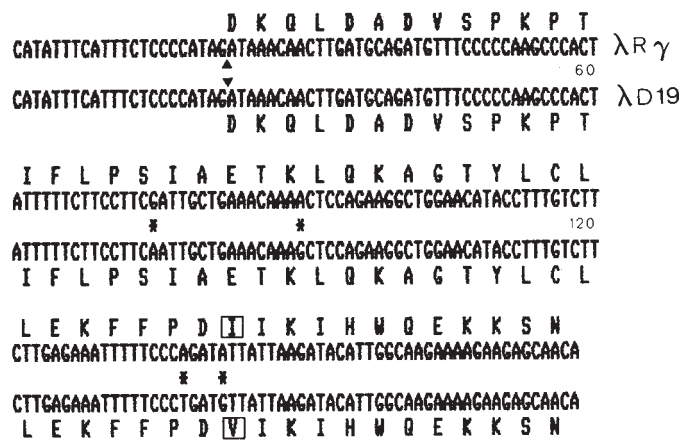


Fig. 3 Allotypic differences in human TRG γ constant-region genes. *Bam*HI and *Hind*III fragments containing TRG γ sequences of λ R γ and λ D19 were cloned in M13 mp9 and sequenced. The sequence shown (upper sequence of each line is λ R γ and lower sequence of each line is λ D19) corresponds to the intron/exon boundary at the start of the TRG γ constant region (the arrows indicate position of the RNA splicing site). The predicted amino-acid sequences start at the codon corresponding to residue 120 of the mouse gene¹⁰. Asterisks between lines indicate base differences and the boxed residues show the allotypic isoleucine-valine substitutions.

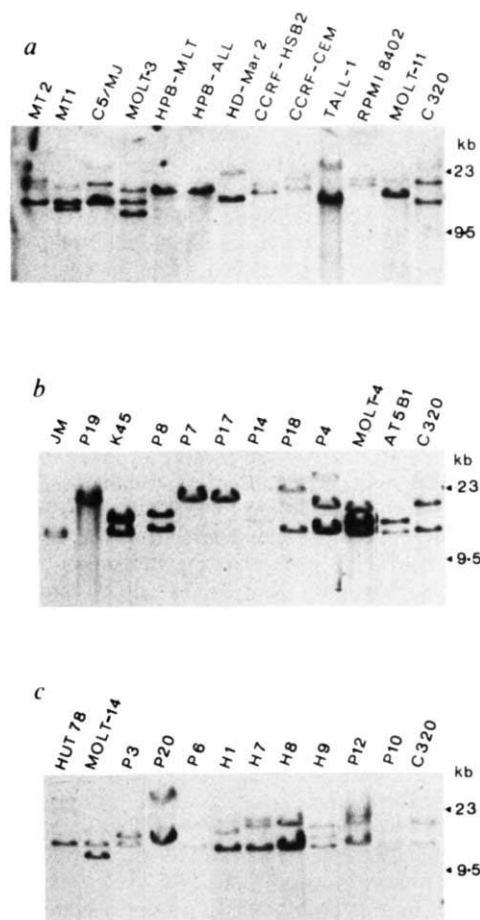


Fig. 4 Filter hybridization of human T-cell DNAs with TRG γ probe M13H60. Genomic DNAs (described in Tables 1 and 2 legends) were digested completely with *Bam*HI, fractionated and transferred to cellulose nitrate. Filters were hybridized with nick-translated M13H60 as described in Fig. 2 legend. Sizes refer to co-electrophoresis of λ phage DNA cut with *Hind*III.

Figure 3 illustrates the homology of nucleotide sequence and derived protein sequence between the two clones in this region of the gene. There are four base differences in this stretch, three being silent changes whereas one is a neutral isoleucine-valine substitution (Fig. 3). Interestingly, this position in the mouse sequence also predicts a valine residue¹⁰. It is probable that the two human TRG γ 1 sequences correspond to the genes encoded by the allelic chromosomes; genomic hybridization data show that Raji DNA is homozygous for allelic chromosome 1 (Fig. 2) whereas the restriction maps of the phage clones obtained from D-PLL show this individual to be heterozygous (Fig. 1). The amino-acid replacement seen between the two genes, therefore, represents a potential allotypic difference in TRG γ 1 gene of human. Our results thus far indicate that human DNA contains only two TRG γ genes while there are three genes in mouse¹², but these genes have not yet been linked in DNA. We cannot completely exclude the presence of more TRG γ genes in humans, particularly as TRG γ 1 and γ 2 genes have very similar restriction maps.

The mouse gene has been shown to undergo rearrangement specifically in T cells, but the rearrangement patterns and gene cloning indicate limited diversity^{11,12}. The situation in human T cells appears to be quite different from mouse with respect to complexity of DNA arrangement observed. The restriction enzyme *Bam*HI digests within TRG γ , yielding two fragments, 15 and 12.5 kb, carrying TRG γ 1 and γ 2 genes, respectively. Preliminary experiments indicate the absence of restriction fragment length polymorphism with this enzyme, which is able to detect rearrangement in T-cell DNA. Analysis of a set of T-cell lines and fresh T-cell leukaemia samples for rearrangement of

Table 1 T-cell lines used in the present studies

Cell line	Surface marker phenotype			
	T3	T4	T8	T11
MT1	-	-	-	+
MT2	-	+	-	-
C5/MJ	-	+	-	+
MOLT-3	-	+	+	+
MOLT-4	-	+	+	+
HPB-MLT	+	+	+	+
HPB-ALL	+	+	+	+
HD-Mar2	+	+	+	+
CCRF-HSB2	+	-	-	-
TALL-1	+	+	+	+
RPM1 8402	±	±	-	+
MOLT-11	-	-	+	+
JM	+	+	+	+
K45	-	-	-	+
HUT78	+	+	-	-
MOLT-14	+	-	-	-

Cell-surface phenotypes and cell lines were obtained from refs 20, 21 and J. Minowada (unpublished).

TRG γ genes revealed, in every case tested, at least one rearranged TRG γ gene. Table 1 lists the T-cell lines and Table 2 the T-cell leukaemia types used. The T-cell line collection includes helper, cytotoxic/suppressor T cells and thymocytes; the T-cell leukaemias include T-cell chronic lymphocytic, acute lymphocytic, prolymphocytic and adult T-cell leukaemia and Sézary syndrome.

Figure 4 shows the results of the DNA rearrangements and compares *Bam*HI-digested T-cell DNAs hybridized with the M13H60 probe with Colo320 DNA (a non-lymphoid DNA¹⁷). The various T-cell samples analysed show a wide range of rearranged restriction fragment sizes. Most DNAs have unique patterns of TRG γ rearrangement; possible exceptions are P7 and P17 (Fig. 4b), H7 which resembles MT2 (Fig. 4a, c) and H9 which resembles the patterns of MOLT-3/MOLT-4 (Fig. 4). This observation indicates a wide variety of rearrangement possibilities which must mean a degree of variability in TRG γ in either V genes, diversity (D) segments and/or joining (J) segments in human DNA. The similar patterns seen in some samples may, however, argue for a relatively restricted V-gene repertoire or even limited rearranged types in leukaemia. (MOLT-4 and MOLT-3 DNA possess an identical TRG γ rearrangement pattern; a similar observation has been made with T-cell receptor β -chain probes (T.H.R., unpublished), and this is because the two cell lines are derived from the same leukaemic patient (J. Minowada, personal communication).)

Apparently both non-allelic TRG γ genes can undergo rearrangement. Further deletion of TRG γ 1 appears to accompany rearrangement to the TRG γ 2 gene. This parallels the human T-cell receptor β -genes where there are also two tandemly organized genes of which the C_{β} 1 gene can be deleted for C_{β} 2 expression^{7,18,19}. Rearrangement of both non-allelic TRG γ genes is illustrated in the DNA of, for example, patient 14; this DNA does not contain TRG γ 1 or γ 2 genes in the unrearranged form but instead has two new, rearranged DNA fragments. The absence of unrearranged TRG γ 2 genes argues that those changes seen are associated with the TRG γ 2 genes on both chromosomes and that TRG γ 1 genes have been deleted. Furthermore, other cells (such as the related HPB-ALL and HPB-MLT) show only one restriction fragment containing TRG γ genes and this is of a new, rearranged size.

It has often been difficult to show the clonality of T-cell leukaemias and also to diagnose lymphomas as B or T cell; T-cell receptor β -chain probes have recently been used to establish T-cell leukaemia clonality¹⁹. The TRG γ gene provides a further potential tool for establishment of the T-cell origin of

Table 2 T-cell leukaemias used

Patient	Diagnosis	Surface phenotype		
		T3	T4	T8
P3	CLL	-		+
P4	Sézary			
P6	Sézary		+	
P7	ALL			
P8	PLL	+	+	-
P10	PLL	+	+	-
P12	PLL	+	-	+
P14	ALL			
P17	ALL			
P18	CLL			
P19	ATL			
P20	ATL			
H1	CLL	+	+	-
H7	Sézary	+	+	-
H8	CLL		+	
H9	CLL		+	

No surface protein data are available for patients 4, 7, 14, 17-20 or AT5B1 in Fig. 4. Patients designated P are described elsewhere¹⁹. Abbreviations: CLL, chronic lymphocytic leukaemia; ALL, acute lymphocytic leukaemia; PLL, prolymphocytic leukaemia; ATL, adult T-cell leukaemia.

leukaemias and lymphomas, as we show that TRG γ gene rearrangement occurs in all T-cell types studied (thymocyte, helper and cytotoxic/suppressor cells). All forms of T-cell leukaemias studied (Table 2) show TRG γ rearrangement, which indicates that the TRG γ gene is an indicator of clonal proliferation characteristic of T-cell leukaemia. The rearrangement of the TRG γ genes seems to parallel that of the T-cell receptor β -chain in its ubiquity of rearrangement in human T cells. This DNA change of the TRG γ genes is therefore presumably an early human T-cell differentiation event.

Finally, the limited number of TRG γ gene rearrangement patterns observed in mouse T cells^{11,12} is apparently quite different from that in human T cells. We show in the latter that great variation occurs between the DNA patterns in a panel of human T cells (Fig. 4). This indicates that, at least in human DNA, significant variability does exist in the TRG γ genes (V and/or D and J) or the regions flanking these genes in different people.

We thank Miss M. A. Stinson and Mr A. Forster for technical assistance, and Drs Catovsky, Luzzatto, Karpas, Minowada, Smith and Hammarstrom for cell lines and leukaemic samples. M.P.L. is the recipient of an EMBO fellowship.

Received 18 March; accepted 14 June 1985.

- Chein, Y.-H. *et al.* *Nature* **312**, 31-35 (1984).
- Chein, Y.-H., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
- Clark, S. P., Yoshikai, Y., Siu, G., Taylor, S., Hood, L. & Mak, T. W. *Nature* **311**, 387-389 (1984).
- Hedrick, S. M., Nielsen, A. E., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
- Saito, H. *et al.* *Nature* **312**, 36-40 (1984).
- Sim, G. K. *et al.* *Nature* **312**, 771-775 (1984).
- Sims, J. E., Tunnacliffe, A., Smith, W. S. & Rabbitts, T. H. *Nature* **312**, 541-545 (1984).
- Siu, G. *et al.* *Cell* **37**, 381-391 (1984).
- Yanagi, Y. *et al.* *Nature* **308**, 145-149 (1984).
- Saito, H. *et al.* *Nature* **309**, 757-762 (1984).
- Kranz, D. M. *et al.* *Nature* **313**, 752-755 (1985).
- Hayday, A. C. *et al.* *Cell* **40**, 259-269 (1985).
- Rabbitts, T. H. *et al.* *EMBO J.* **4**, 1461-1465 (1985).
- Snodgrass, H. R., Dembic, Z., Steinmetz, M. & von Boehmer, H. *Nature* **315**, 232-233 (1985).
- Hamlyn, P. H. & Rabbitts, T. H. *Nature* **304**, 135-139 (1983).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
- Altalio, K., Schwab, M., Lin, C. C., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1707-1711 (1983).
- Toyonaga, B., Yanagi, Y., Suciu-Foca, N., Minden, M. & Mak, T. W. *Nature* **311**, 385-387 (1984).
- Rabbitts, T. H. *et al.* *EMBO J.* (in the press).
- Minowada, J. *et al.* *Current Concepts in Human Immunobiology* (eds Serrou, B. *et al.*) 75-84 (Elsevier, Amsterdam, 1982).
- Greaves, M. F. *et al.* *Leukaemia Res.* **5**, 281-299 (1981).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
- Bernheim, A., Berger, R. & Lenoir, G. *Cancer Genet. Cytogenet.* **3**, 307-315 (1981).
- Klein, E. *et al.* *Cancer Res.* **28**, 1300-1310 (1968).

Detergent, Teflon tubing and specs

This month's batch of new products includes a left-handed glove box, stereo magnifying spectacles, Teflon tubing the diameter of human hair and a new version of Triton X-100 detergent.

- A new, cloned, FPLCpure Klenow fragment of DNA polymerase I is now available from Pharmacia's Molecular Biology Division for researchers conducting dideoxy sequencing. Unlike traditional preparations in which (after proteolytic digestion) the large Klenow fragment must be separated from other over- and under-digested fragments, the over-producing cloned strain allows for production of a more homogeneous product. The Pharmacia preparation is purified by FPLC liquid chromatography at the final step, ensuring that the new FPLCpure Klenow fragment has the highest functional purity available.

Reader Service No. 100.

- A new kit from Anglian Biotechnology provides all the materials needed for oligonucleotide site-directed mutagenesis in M13 phage. The use of the single-stranded DNA M13 phage is ideal for mutagenesis with synthetic oligonucleotides and for sequencing DNA using the dideoxy method. Since precise changes can be made in any DNA fragments cloned into M13 using synthetic oligonucleotides, a defined system can provide a simple way of generating this mutation. Anglian Biotechnology is marketing a comprehensive protocol and experimental manual written by Carter, Bedouelle, Waye and Winter, of the MRC Laboratory of Molecular Biology in Cambridge, and a site-directed mutagenesis kit.

Reader Service No. 101.



New Dewar flasks from Pope.

- Newly designed Low-Form Shallow Dewars from Pope Scientific Inc. permit operations not practical with standard cylindrical or spherical Dewars. These operations include rapid freezing and recovery of tissue/organ specimens; insertion of small flasks in the Dewar; agitation in the flask with a magnetic stirrer; and tempering of stirring rods, pestles and other small objects. The cooling range is from dry ice/acetone to liquid nitrogen temperatures. The flasks are available in two configurations, hemispherical or cylindrical.

Reader Service No. 102.

These notes are based on information provided by the manufacturers. To obtain further details about these products use the reader service card bound inside the journal.

- A.R.B. Computer Consultants' new Integrated Word Processing Systems are developed from A.R.B.'s Scientific Word Processing System released last year. There are two models in the new series; the fully integrated model with six full character sets and the scientific model with only three. They feature Tanberg 2200s terminals and a S900 letter quality vario-printer. Both models have English, Greek and scientific characters which may be displayed in either superscript or subscript form. All character sets may be displayed or printed on the screen normal, bold or underlined. The integrated model can display forms and tables, 41 languages and a range of chemical pictures.

Reader Service No. 103.

- Ligand binding analysis programs for saturation and displacement binding assays are now available from Lundon Software. These programs feature ease of data entry and correction, graphical video display, printer graphics, statistical evaluation and highly efficient computation of parameter values. Programs are available for the IBM PC series, Apple II series, and the DEC-VAX.

Reader Service No. 104.

- The West German company Krombach & Sohn has developed a new range of stereo magnifying spectacles. The multiple coating and high colour correction of the wide field optics in the spectacles give extremely high resolution and depth of field. There is a choice of magnifications from 2.2 $\times$ to 10 $\times$ with extremely long working distances, up to 420 mm. The spectacles are adjustable for interpupillary distance and convergence, and each side is independently focusable to allow for differences in the user's eyes.

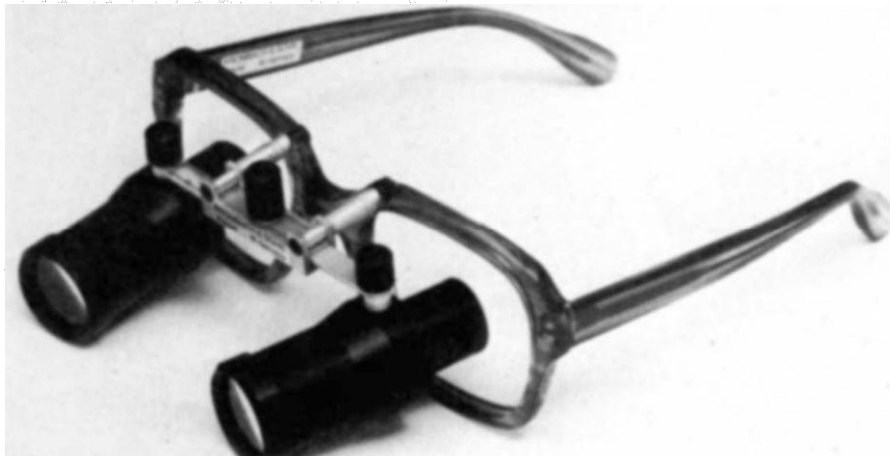
Reader Service No. 105.



A safer bottle has been invented by May & Baker.

- May & Baker aims to improve the safety of dispensing and handling hazardous liquid chemicals in the laboratory with the introduction of a new range of specially designed glass bottles. Unlike traditional narrow-mouthed 'Winchester' glassware, the May & Baker bottles incorporate features designed to reduce the risk of wasteful and potentially dangerous liquid-splash. The wider mouth of the bottles, together with their newly developed shoulder and base profiling, gives a large reduction in the turbulence and 'glugging' experienced when dispensing liquids. The bottles have a redesigned mouth-lip which virtually eliminates the formation of drips. And for corrosive products, the bottles are additionally fitted with a plastic anti-drip collar which gives an even cleaner break of liquid flow.

Reader Service No. 106.



Krombach's stereo magnifying spectacles are capable of magnifications varying between 2.2 $\times$ and 10 $\times$.

• Latest additions to Bio-Yeda's range of monoclonal and polyclonal cytoskeletal antisera include anti-desmin (rabbit), monoclonal anti-vinculin, anti-alpha-actinin (rabbit) and anti-chicken spectrin (rabbit). Antibody titre performance is determined by immunofluorescent labelling of cultured cells and frozen tissue sections and by immunoperoxidase staining of paraffin-embedded formalin-fixed tissue sections. Antibody specificity of each lot is tested by immunoblotting. Bio-Yeda's antisera to cytoskeletal proteins show high cross-reactivity with corresponding proteins in other mammalian and many non-mammalian species.
Reader Service No. 107.

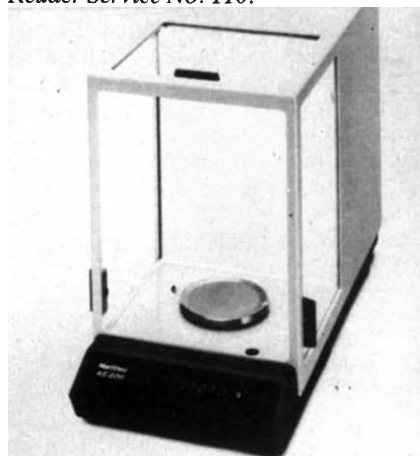


Indirect immunofluorescence labelling of chicken fibroblast cells.

• Two new products have been added to the Cambridge Research Biochemicals range of peptide antisera for radioimmunoassay: atrial natriuretic peptide antiserum and enkephalin antiserum for the specific measurement of total enkephalins and/or [Leu⁵] enkephalin, and/or [Met⁵] enkephalin.
Reader Service No. 108.

• New items in the latest Medix Biochemica catalogue of monoclonal antibodies include monoclonal anti-human prolactin, anti-human growth hormone and anti-human carcino-embryonic antigen. As well as supplying 'laboratory' quantities of monoclonals, Medix Biochemica produces monoclonal antibodies on an industrial scale.
Reader Service No. 109.

• Equipped with dual refractive index and UV detectors, Bio-Rad's organic acids HPLC analysis system is able to accurately monitor fermentable carbohydrates and the major by-products of fermentation. At the start of fermentation carbohydrates are profiled. As the process progresses, carbohydrate consumption and subsequent alcohol and by-product appearance are easily measured. A refractometer detects fermenting carbohydrates, glycerol and alcohol, while a UV monitor, at 210 nm, simultaneously detects volatile fatty acids and other fermentation by-products.
Reader Service No. 110.



The Mettler AE balance.

• The Mettler AE series of analytical balances has been expanded with the introduction of the AE200. This new analytical balance, with 0...205 g weighing range and a readability 0.1mg, allows precise dispensing into standard as well as heavier containers. Heavier containers can also be used to perform back-weighings successfully. Various options make it possible to transfer data from the balance via RS232C, CL and IEEE 488 interfaces. The balance is operated with a single control bar and the weighing chamber is freely accessible from three sides. An adjustable stability detector and a selectable integration time extension make the AE200 analytical balance adaptable to specific operating conditions.
Reader Service No. 111.

How to avoid the 19 critical mistakes in buying fermentation equipment

Make one mistake and you can have problems that include inadequate process control, permanent operator inconvenience and costly downtime.

Now NBS has published a comprehensive review of the problems and pitfalls that await the unwary buyer of industrial fermentors — and some very practical suggestions to help avoid them.

Called *19 Critical Mistakes Made in Buying Fermentation Equipment and What You Can Do to Avoid Them*, this practical guide is available without charge.

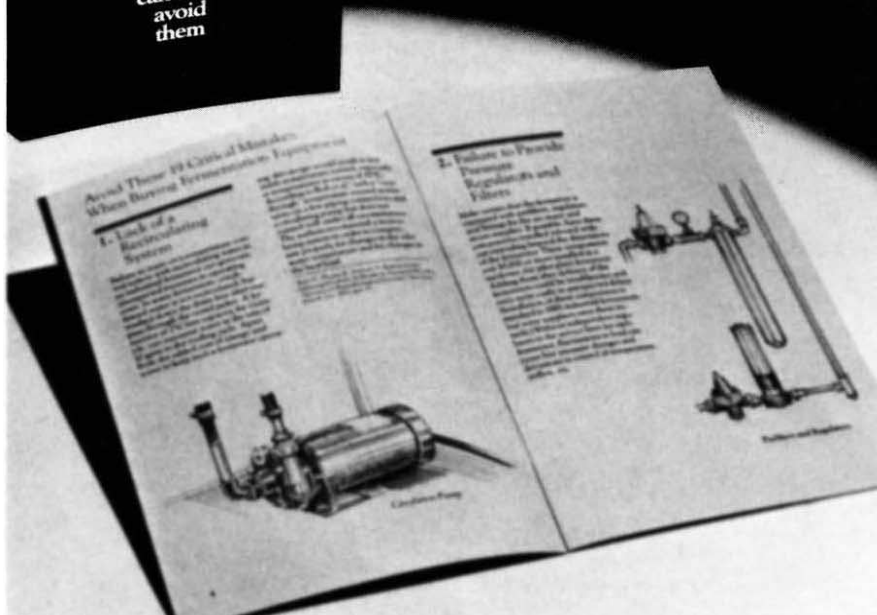
Send for your free copy today. New Brunswick Scientific Co., Inc., P.O. Box 986, Edison, NJ 08818. (201) 287-1200 or (800) 631-5417.

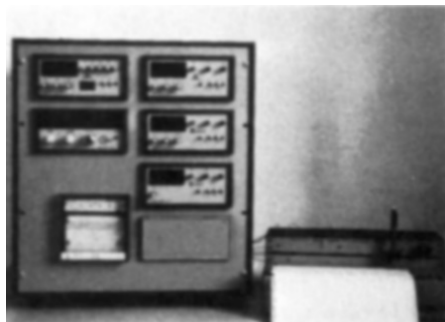


NEW BRUNSWICK SCIENTIFIC

At NBS Advanced Technology is a Way of Life

**19
CRITICAL
MISTAKES**
made in buying
fermentation
equipment
and
what you
can do to
avoid
them





Fermentation control in easy stages by Biologic.

• A modular system for the computer-based control of biotechnology-based processes has been developed by Biologic of Versailles, France. Laboratory, pilot plant and small-scale production units can be controlled with this package. Unit 1 is a modular analogic cabinet; Unit 2 (Bio-Stel) is an automaton (128 I/O) oriented towards sterilization sequences; Unit 3 (Bio-Mak) is an acquisition and/or command microprocessor-based system which can also be operated with an external computer, via RS232C or RS422/423 or IEEE data link; Unit 4 (Bio-Net) is an industrial local area network offering low cost, low speed (50 Kbps), 256 logical nodes, 4 km long, and allowing for the connection of different types of equipment through suitable connectors.

Reader Service No. 112.



The Schleicher & Schuell Bio-Trap.

• A new inert membrane from Schleicher & Schuell retains charged macromolecules in an electric field without adsorption. This membrane retains macromolecules larger than 5,000 daltons in size. Application of the inert membrane in the Bio-Trap dialysis system enables the user to elute charged macromolecules (DNA, RNA and proteins) from gels, to concentrate, dialyse and separate them isoelectrically. In the Bio-Trap, the macromolecules are driven into a 'membrane-trap' by means of an electric field where they are quantitatively caught in a volume of 200 to 400 microlitres and are easily removed with a Pasteur pipette. The recovered macromolecules are suitable for enzyme digestion or other experiments without centrifugation, precipitation or other steps. Bio-Trap can be employed without additional equipment in a conventional electrophoresis chamber.

Reader Service No. 113.

• A new high-intensity light point source incorporating a mercury arc lamp is being manufactured by the Hampshire-based company Pulse Photonics Ltd. The system is available with 50W, 100W and 200W power sources, and has applications for Schlieren photography, holography and microscopy, as well as many uses in the research/industrial fields. The system has a high-voltage power supply for starting purposes, featuring an automatic circuit design, making it virtually unaffected by lead length. The lamp housing incorporates an easy-to-focus lens.

Reader Service No. 114.



A left-handed glove box from Labconco.

• For labs with limited space for a glove box, a new option of 'left-handedness' from Labconco could be the answer. Labconco offers four glove boxes—all of which are now available in both the left-handed and right-handed position. The Radioisotope Glove Box is designed for low-level radioactive procedures. For general-purpose controlled atmosphere work, the Controlled Atmosphere Glove Box is available. The Bacteriological Glove Box (which has an access door but no interchange chamber) is designed to handle a wide range of biological material. Finally, the Chemical Carcinogen Glove Box, equipped with an activated adsorber, is ideal for the containment of potentially carcinogenic research materials.

Reader Service No. 115.

• Boehringer Mannheim Biochemicals has introduced a specially purified preparation of Triton X-100. Previously available commercial preparations of this detergent have contained high levels of peroxides, which can cause artefacts in the isolation of proteins, particularly in their tertiary structure. They may also interfere with any assay which depends on the formation or consumption of reducing reagents.

Reader Service No. 116.

• The EC910 transmission densitometer from EC Apparatus, interfaced with an Apple II Plus or IIe computer, provides a video display of pattern, integration of peak areas, printout of optical density trace and integration as well as data storage on disk. The EC910 has fine-line resolution capabilities in the optical density range 0-4 with zero background suppression using apertures from a 0.05-mm point to a 1.3-mm slit, complete resolution of 100- μ m spaced bands, and visible range selection via optical filters.

Reader Service No. 117.

©1985 Nature Publishing Group

ADVERTISEMENTS

**PHORBOL
ESTERS**

**TOOLS FOR CELL
BIOLOGY AND
CANCER RESEARCH**



LC SERVICES CORPORATION

165 New Boston Street
Woburn, MA 01801
(617) 938-1700

Reader Service No.3

LIPOSOMES

**REPRODUCIBLE - HOMOGENEOUS -
EXTREMELY RAPID**

Our new commercial LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized

- UNILAMELLAR OR
- OLIGOLAMELLAR OR EVEN
- MULTILAMELLAR LIPOSOMES

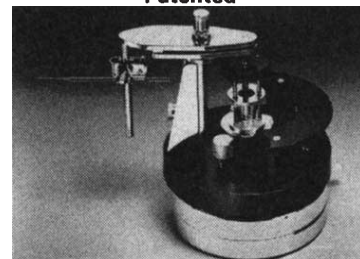
depending on the used experimental conditions. Liposome size is selectable between 25 and ca. 600 nm diameter up to 1000 nm for multilamellar liposomes. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. High encapsulation efficiency for hydrophilic and/or lipophilic agents. Lipid concentration can be up to 300 mg/ml. Lipid loss during preparation is extremely low: down to 0.1%.

DIANORM - Geräte

POB 126, D-8000 Munich 65, FRG

Reader Service No.4

**SINGER
MICROMANIPULATOR
MK III
Patented**



Perfect control in 3 Dimensions from a single joystick. Continuous variable reductions to suit all magnifications. No after sales maintenance needed.

Details from:
Singer Instrument Co. Ltd.,
Treborough Lodge Roadwater,
Watchet Somerset TA23 0QL
England. Tel. (0984) 40226 Telex:
337492 Comcab G.

Reader Service No.7

NEWAT '85

NEWDAT data base now available on 5¼" floppy disks for IBM PC's and look-alikes. More than 1200 amino acid sequences with literature citations. Six-disk kit includes programs for entering data, searching and aligning. \$250. Limited supply.

Telephone: 619-453-2467

or

**NEWAT
P.O. Box 12822,
La Jolla
CA 92037 USA**

Reader Service No.2



Zeus Teflon tubing is light-conducting, ultra-thin and impervious to most solvents.

• A capillary tubing made of Teflon for electronic and computer applications has been released by Zeus. The new tubing, Zeus Sub-Lite-Wall, features superior lubricity and is impervious to virtually all solvents and chemicals. Operating temperatures range from cryogenics to +500°F. Internal diameters smaller than a human hair are available with metric uniformity and tolerances. Samples are available from Zeus, based in Holzkirchen, West Germany. *Reader Service No. 118.*



Autosorb-6 from Quantachrome.

• An advanced system for the multiple analysis of powders has been developed by Quantachrome Corporation, and it is now available in the UK from Techmation. The Autosorb-6 is capable of simultaneously measuring surface area, pore size distribution of up to six samples in each loading. Full data handling and plotting facilities are provided, including multipoint adsorption/desorption isotherms, Langmuir, BJH and single and multipoint BET methods. The software can also analyse pore radius, pore radius distribution and total pore volume. The computerized Autosorb-6 incorporates full data handling and plotting facilities. *Reader Service No. 119.*

• Packard Instrument Company's new computer-aided liquid scintillation analyser, the Tri-Carb 2000CA, applies new techniques to liquid scintillation counting (LSC) such as spectrum unfolding and three-dimensional spectrum analysis to enhance count results of LSC samples. Spectrum unfolding takes the composite spectrum of dual labelled samples and "unfolds" each spectrum for accurate c.p.m. and d.p.m. calculation of the individual radionuclides. Three-dimensional spectrum analysis differentiates background radiation from beta decay events and gives the system high counting performance and low-level counting sensitivity. Packard uses computer-aided liquid scintillation analysis (CALSA), a 'knowledge-based' computer technique which provides access to high-level languages and powerful applications programs. The Tri-Carb 2000CA also supports user-written routines, in which LSC data can be custom formatted into reports or merged with other laboratory data. CALSA is based on the IBM-PC microcomputer. *Reader Service No. 120.*



The Packard Tri-Carb 2000CA for liquid scintillation counting.

• The Perkin-Elmer LIMS/2000 laboratory information management system is designed to be easily customized to meet a laboratory's particular requirements. Running on standard Perkin-Elmer computer hardware, LIMS/2000 can be used in three ways: as a dedicated information management system; as a tool to provide automation of the instruments in a laboratory; or as a fully integrated instrument automation and information management system. LIMS/2000 can track samples through the laboratory; assign test requirements to each sample at the point of log-in; produce worklists and backlog reports; collect test results automatically from instruments (of any manufacture) and take data manually from terminals; generate a final report for each sample by collating all the test results; archive data for long-term storage; and search this database according to any chosen criteria. Other features include: a queue manager for easy rescheduling of samples; interactive colour graphics; and comprehensive automation software. *Reader Service No. 121.*



A pilot-scale cell levitation fermentor fully instrumented with Chemap AG systems for fermentation monitoring. The fermenting set-up is suitable for growing lymphoblasts, hybridomas and various mammalian cells. Use the reader service card (no. 122) for further details.

• The calibration of automated Coulter counters can be verified in minutes using the new Coulter S-Cal calibrator kits. Calibrator values for WBC, RBC, MCV and Plt parameters are derived by replicate analyses of fresh whole blood samples on several correctly calibrated Coulter S or S-plus Series instruments. Those assigned values are traceable to reference methods with an accuracy within $\pm 2\%$. *Reader Service No. 123.*

• Jansen's new biochemical L-*p*-bromotetramisole (R 30 402) is more potent than levamisole as an inhibitor of alkaline phosphatase. The inactive *d*-isomer *d*-*p*-bromotetramisole (R 30 401), is useful as an internal control. The organo and stereospecificities of these alkaline phosphatase inhibitors have been demonstrated biochemically as well as cytochemically in a variety of tissues and species. *Reader Service No. 124.*

• Chemical Dynamics Corporation has added atrial natriuretic peptides to its line of biologically active peptides. For a free copy of the special peptide issue of Chemical Dynamics' publication *Chemalog Hi-Lites* describing atrial natriuretic peptides, use the Reader Service Card. *Reader Service No. 125.*

• A new development, from Accurate Chemical, Accu-rett tubes are a convenient, reliable method of counting reticulocytes by the improved Romanosky staining method using light microscopy. Featuring long-life and room temperature stability, each tube is already to use, and contains a precise amount of dried stain. Accu-rett tubes eliminate in-house batch variability, and the time-consuming process of preparing stain solution. *Reader Service No. 126.*